

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 03:35:56 / Search time 5459 Seconds  
(without alignments)  
11579.281 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 2172  
Sequence: 1 atgsgggagatgaagatcgga.....tagatgagatgcatggaat 2172

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 segs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: GenEmbl:\*  
2: gb\_ba:\*  
3: gb\_htg:\*  
4: gb\_in:\*  
5: gb\_om:\*  
6: gb\_ov:\*  
7: gb\_pat:\*  
8: gb\_ph:\*  
9: gb\_pl:\*  
10: gb\_pr:\*  
11: gb\_ro:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pla:\*  
35: em\_htg\_ror:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vit:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2172	100.0	2224	9 AY004154	AY004154 Homo sapi
2	2164	99.6	2211	9 AF262962	AF262962 Homo sapi
3	1900.6	87.5	2189	9 HSM802178	AL137462 Homo sapi
4	1577.6	72.6	3030	10 S75275	S75275 RVLG=vasa-1
5	1562	71.9	1984	9 AK093439	AK093439 Homo sapi
6	1443	66.4	1930	10 MUSDVH	D14859 Mouse mRNA
7	751.2	34.6	1989	5 AB004836	AB004836 Gallus ga
8	712	32.8	2502	5 AF046043	AF046043 Xenopus 1
9	686.8	31.6	2865	5 AB005147	AB005147 Dario rer
10	686.4	31.6	2492	5 DRY12007	Y12007 D. rerio vlg
11	619.4	28.5	2058	5 AB051835	AB051835 Oreochrom
12	619.4	28.5	2130	5 AB032467	AB032467 Oreochrom
13	612.2	28.2	2612	5 AB032566	AB032566 Oncorhynch
14	565	26.0	1512	5 AF479823	AF479823 Pantodon
15	552.8	25.5	2187	5 AB063484	AB063484 Oryzias 1
16	548.2	25.2	1158	5 AF251800	AF251800 Dario dan
17	543.8	25.0	1730	5 AF479821	AF479821 Hyphessob
18	511.8	23.6	1972	5 AF479824	AF479824 Cyprinus
19	503.4	23.2	1506	5 AF479824	AF479824 Melanotae
20	499.6	23.0	1586	5 AF520608	AF520608 Sparus au
21	498.6	23.0	1816	5 AF479825	AF479825 Oncorhynch
22	487.4	22.4	2551	5 AB016603	AB016603 Ciona int
23	487.4	22.4	1873	3 AB016604	AB016604 Ciona int
24	457	21.0	1873	3 AB047385	AB047385 Ephedratia
25	456.6	21.0	3132	3 AB047383	AB047383 Hydra mag
26	438	20.2	3023	3 AB047803	AB047803 Ciona sav
27	436.4	20.1	2777	3 AB047802	AB047802 Ciona sav
28	434.6	20.0	2721	3 AB047382	AB047382 Hydra mag
29	428.8	19.7	1449	5 AF479822	AF479822 Oryzias 1
30	383.2	17.6	2319	9 HSAF000985	AF000985 Homo sapi
31	383.2	17.6	4416	9 HSAF000984	AF000984 Homo sapi
32	378.2	17.4	2384	3 AB047381	AB047381 Hydra mag
33	377	17.4	524	4 AY100475	AY100475 Equus cab
34	371	17.1	2187	6 AX305316	AX305316 Sequence
35	371	17.1	2187	10 MUSRNAHELI	L25126 Mus musculu
36	371	17.1	2201	9 AF061337	AF061337 Homo sapi
37	371	17.1	2508	9 BC011819	BC011819 Homo sapi
38	371	17.1	3158	9 HSU50553	U50553 Homo sapien
39	371	17.1	3185	10 MMBRNRAHL	Z38117 M. musculus
40	370.2	17.0	3731	10 MMU7376	AJ007376 Mus muscu
41	370.2	17.0	3767	10 BC021453	BC021453 Mus muscu
42	369.4	17.0	1989	10 AB080116	AB080116 Mesocric
43	369.4	17.0	3408	9 HSAF000983	AF000983 Homo sapi
44	369.4	17.0	5332	9 HSAF000982	AF000982 Homo sapi
45	358	16.5	1983	10 MUSPL10A	J04847 Mouse PL10

## ALIGNMENTS

RESULT 1  
AY004154  
LOCUS  
DEFINITION Homo sapiens DEAD box RNA helicase (VASA) mRNA, complete cds.  
ACCESSION AY004154  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 2224)  
Castillon,D.H., Quade,B.J., Wang,T.Y., Quigley,C. and Crum,C.P.  
The human VASA gene is specifically expressed in the germ cell lineage

Pred. No. is the number of results predicted by chance to have a

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9585-9590 (2000)  
 MEDLINE 20402578  
 PUBMED 10920202  
 REFERENCE 2 (bases 1 to 2224)  
 AUTHORS Castriillon, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-JUL-2000) Pathology, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115, USA  
 FEATURES  
 SOURCE Location/Qualifiers  
 1. 2224  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /map="5q11.2-12"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 1. 2224  
 /gene="VASA"  
 14. 2188  
 /gene="VASA"  
 /codon\_start=1  
 /product="DEAD box RNA helicase"  
 /protein\_id="AA086585.1"  
 /db\_xref="GI:9438227"  
 /translation="MGDEDEAEINPMSYVPIFEKORYSGENDNPNRTPASSSEM  
 DDGSRBDHEPKSGFASGRNFGNDAGCNDRNTSMGGVGKSGNGGFSNRE  
 DDDSSGFRBSSNDCEDNPTNRGPRKRGTRDGNSEASPRPYRGGRGSRGCRGPF  
 GLGSPNDLDBECMQRTGLFGRPRPLVSTGNGDVSQSSGSGSEGGKGNBEV  
 ITGSKMSKSEAEAGSGSDYQPRKVIYIPPEDEDSIFAHYQTGIPDKYDILV  
 EVSGHDAEPALITFEENALCQTLNNNTAKAYETLTPVKXSIP11ILAGRDLMACAT  
 GSGKTAFLPLILAHMHMDGITASRFEKLOPECIIVAPTELVNOQIYLKRFSGT  
 CVRAVLYGTGOLGHSIROIVOGNIIICATGRMLDIIGKEIGLKOIKYLVLEADR  
 MLMGDFEEMKLLSCPGMRBOROTLMSATPEPILORALFSLSNYIFVAVGV  
 GGAARDVOQTVLYQGFQSKREKLVEILRNIDERTMFWETKKAADPTATLCEKXS  
 TTSIHGREOREQALGDFRFGKCPVLVATSVARGLDIENVQVINFPLSTIDEX  
 VHRIGRGRCGNTRAISFPDLESNDMLAQPLVLTDAQODVPMLEELAFSTYIIG  
 FSGTRGVNFVSVDTRKSKSTLNTAGFSSSRAPNVDDESMD"

gene  
 CDS  
 14. 2188  
 /gene="VASA"  
 /codon\_start=1  
 /product="DEAD box RNA helicase"  
 /protein\_id="AA086585.1"  
 /db\_xref="GI:9438227"  
 /translation="MGDEDEAEINPMSYVPIFEKORYSGENDNPNRTPASSSEM  
 DDGSRBDHEPKSGFASGRNFGNDAGCNDRNTSMGGVGKSGNGGFSNRE  
 DDDSSGFRBSSNDCEDNPTNRGPRKRGTRDGNSEASPRPYRGGRGSRGCRGPF  
 GLGSPNDLDBECMQRTGLFGRPRPLVSTGNGDVSQSSGSGSEGGKGNBEV  
 ITGSKMSKSEAEAGSGSDYQPRKVIYIPPEDEDSIFAHYQTGIPDKYDILV  
 EVSGHDAEPALITFEENALCQTLNNNTAKAYETLTPVKXSIP11ILAGRDLMACAT  
 GSGKTAFLPLILAHMHMDGITASRFEKLOPECIIVAPTELVNOQIYLKRFSGT  
 CVRAVLYGTGOLGHSIROIVOGNIIICATGRMLDIIGKEIGLKOIKYLVLEADR  
 MLMGDFEEMKLLSCPGMRBOROTLMSATPEPILORALFSLSNYIFVAVGV  
 GGAARDVOQTVLYQGFQSKREKLVEILRNIDERTMFWETKKAADPTATLCEKXS  
 TTSIHGREOREQALGDFRFGKCPVLVATSVARGLDIENVQVINFPLSTIDEX  
 VHRIGRGRCGNTRAISFPDLESNDMLAQPLVLTDAQODVPMLEELAFSTYIIG  
 FSGTRGVNFVSVDTRKSKSTLNTAGFSSSRAPNVDDESMD"

BASE COUNT 678 a 385 c 550 g 611 t

Query Match 100.0%; Score 2172; DB 9; Length 2224;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGGATGAAGATTGGGAACGAAATCAACCTCATATGCTTCTATGTTCCCTA 60  
 14 ATGGGGATGAAGATTGGGAACGAAATCAACCTCATATGCTTCTATGTTCCCTA 73  
 61 TTTGAGAAGATAGTATTCTGAGAAATAGAGACATTTTAAACAGAGCTCCAGCTTCA 120  
 74 TTTGAGAAGATAGTATTCTGAGAAATAGAGACATTTTAAACAGAGCTCCAGCTTCA 133  
 121 TCATCAGAAATAGATGATGACCTTCTCGAAGAGATCATTTTATGAAAAAGTGATTTGCC 180  
 134 TCATCAGAAATAGATGATGACCTTCTCGAAGAGATCATTTTATGAAAAAGTGATTTGCC 193  
 181 TCTGGCCGGAATTTTGGAAAACAGAGATGCTGTGATGATTAACCGGAGATTAATCAATCC 240  
 194 TCTGGCCGGAATTTTGGAAAACAGAGATGCTGTGATGATTAACCGGAGATTAATCAATCC 253  
 241 ACAATGAGTGTGTTTGGAGTTGAAAAGATTGGAAAACAGAGTTTTCCAAACAGCAGG 300  
 254 ACAATGAGTGTGTTTGGAGTTGAAAAGATTGGAAAACAGAGTTTTCCAAACAGCAGG 313  
 301 TTTGAGAAGTGTAGTATGCTGTGTTTCTGAGAGAGTCTAGTAATGACTGCGAAGATTAAT 360  
 314 TTTGAGAAGTGTAGTATGCTGTGTTTCTGAGAGAGTCTAGTAATGACTGCGAAGATTAAT 373  
 361 CCAACACGGAACAGAGGTTTTCACAGAGAGCGGCTATCGAGATGGAATTAATTCACAA 420  
 374 CCAACACGGAACAGAGGTTTTCACAGAGAGCGGCTATCGAGATGGAATTAATTCACAA 433  
 421 GCTTCAGGCAATACAGAAAGGTGGAAGAGTATGTTCCGAGGTGTCGCGTGAAGATTT 480

434 GCTTCAGGCAATACAGAAAGGTGGAAGAGTATGTTCCGAGGTGTCGCGTGAAGATTT 493  
 481 GGTCTAGGAAGTCCAAATTAATGACTTAGACCCACAGAGATGATGACGCCACTGCTGCG 540  
 494 GGTCTAGGAAGTCCAAATTAATGACTTAGACCCACAGAGATGATGACGCCACTGCTGCG 553  
 541 CTTTGGTCTCTGAAAGACAGATTAATGAGGACAGGTAATGATGATCTTCTCAAGC 600  
 554 CTTTGGTCTCTGAAAGACAGATTAATGAGGACAGGTAATGATGATCTTCTCAAGC 613  
 601 AGAAGTGGCAGTGAAGTGAACAGAGTGTGTTACAAAGTITTAATGAAGATTAATACA 660  
 614 AGAAGTGGCAGTGAAGTGAACAGAGTGTGTTACAAAGTITTAATGAAGATTAATACA 673  
 661 GGTCTGGAAGAAATTTTGGAAAGTCAAGACGAGAGAGAGAGAAATAGTATCTCA 720  
 674 GGTCTGGAAGAAATTTTGGAAAGTCAAGACGAGAGAGAGAGAAATAGTATCTCA 733  
 721 GGAACCAAAAGTGAACCTCATATCCCTCTCCACCTGAGATGAGACTCATCTTTGCA 780  
 734 GGAACCAAAAGTGAACCTCATATCCCTCTCCACCTGAGATGAGACTCATCTTTGCA 793  
 781 CATTATCAGACAGGATTAATCTTGAACAAATACAGACTATTTCTTGGAAAGTCTTGA 840  
 794 CATTATCAGACAGGATTAATCTTGAACAAATACAGACTATTTCTTGGAAAGTCTTGA 853  
 841 CATGATCACCACCAAGCAATTTCTGACTTTTGAAGAAAGTAAATCTCTGACAGACTGAAT 900  
 854 CATGATCACCACCAAGCAATTTCTGACTTTTGAAGAAAGTAAATCTCTGACAGACTGAAT 913  
 901 AACCAACTGCTAAAGTGTGTTACTTAAGCTTACTCTGTGCAAAATACAGATTTCT 960  
 914 AACCAACTGCTAAAGTGTGTTACTTAAGCTTACTCTGTGCAAAATACAGATTTCT 973  
 961 ATCATACTTGCAGACAGAGATTTGATGCTTGTGCTCAACAGGCTTGGGAAGACTGCG 1020  
 974 ATCATACTTGCAGACAGAGATTTGATGCTTGTGCTCAACAGGCTTGGGAAGACTGCG 1033  
 1021 GCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGATTAATCTGCGACCTGTTT 1080  
 1034 GCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGATTAATCTGCGACCTGTTT 1093  
 1081 AAGAGTTGAGAGACAGAGATGATTAATGATGATGATGATTAATCTGCGACCTGTTT 1140  
 1094 AAGAGTTGAGAGACAGAGATGATTAATGATGATGATGATTAATCTGCGACCTGTTT 1153  
 1141 ATTATTTGAGAGACAGAGATTTTCTTTGGGACTTGTGTAAGCTGTTGTTATAT 1200  
 1154 ATTATTTGAGAGACAGAGATTTTCTTTGGGACTTGTGTAAGCTGTTGTTATAT 1213  
 1201 GGGGGAACCCAGCTGGGACATTTCAATTCGACAAATTAATGATTAATGATTAATGAT 1260  
 1214 GGGGGAACCCAGCTGGGACATTTCAATTCGACAAATTAATGATTAATGATTAATGAT 1273  
 1261 GCTACTCTGGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 1274 GCTACTCTGGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1333  
 1321 AATATCTTGAAGTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
 1334 AATATCTTGAAGTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1393  
 1381 AAGAGTGAATTTTCTGCGACAGAAATGCAATCAAGAGACAGGCCCAACCTTATATGTC 1440  
 1394 AAGAGTGAATTTTCTGCGACAGAAATGCAATCAAGAGACAGGCCCAACCTTATATGTC 1453  
 1441 AGTCAACTTTTCCAGAGAAATCAAGAGTGTGCTCAGAGTGTGTTTAAAGTCAATTAAT 1500  
 1454 AGTCAACTTTTCCAGAGAAATCAAGAGTGTGCTCAGAGTGTGTTTAAAGTCAATTAAT 1513  
 1501 CTGTTTGTGCTGTTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560

Db 1514 CTGTTTGTGCTGTGGCAAGTGGGTGGAGCATGTGAGATGTTTCAGCAGACCGTTCTC 1573  
 QY 1561 CAAGTTGGCCAGTTCCTCAAAAAGAGAAAAGCTCGTTGAAATTCGCAAAACATAGGGGAT 1620  
 Db 1574 CAAGTTGGCCAGTTCCTCAAAAAGAGAAAAGCTCGTTGAAATTCGCAAAACATAGGGGAT 1633  
 QY 1621 GAAAGAACTATGCTGTTGTTGAAACTAGAGAAAAGCAGATTTTACTGCAATTTTCTT 1680  
 Db 1634 GAAAGAACTATGCTGTTGTTGAAACTAGAGAAAAGCAGATTTTACTGCAATTTTCTT 1693  
 QY 1681 TGTCAAGAAAAAATATCACTACAGATATCCATGATGATCGGGAACAGAGAGACGGGAG 1740  
 Db 1694 TGTCAAGAAAAAATATCACTACAGATATCCATGATGATCGGGAACAGAGAGACGGGAG 1753  
 QY 1741 CAAGCTCTTGAGAGATTTTGGCTTTGAAAGAGCCAGTCTTGTGCTACTTCACTAGAGCT 1800  
 Db 1754 CAAGCTCTTGAGAGATTTTGGCTTTGAAAGAGCCAGTCTTGTGCTACTTCACTAGAGCT 1813  
 QY 1801 GCCAGAGGGCTGATATTTGAAATGTGCAACATGTTATCAATTTTGAATCTTCTTCAAC 1860  
 Db 1814 GCCAGAGGGCTGATATTTGAAATGTGCAACATGTTATCAATTTTGAATCTTCTTCAAC 1873  
 QY 1861 ATTGATGAATATATGTTCAATCGAATTTGGGCGTACTGCTGTTGGGAATACGCGAGACA 1920  
 Db 1874 ATTGATGAATATATGTTCAATCGAATTTGGGCGTACTGCTGTTGGGAATACGCGAGACA 1933  
 QY 1921 ATTTCCTTTTGTGATCTTGAATCGGATACCATTTAGACAGCCCTAGTAAAGTATTTG 1980  
 Db 1934 ATTTCCTTTTGTGATCTTGAATCGGATACCATTTAGACAGCCCTAGTAAAGTATTTG 1993  
 QY 1981 ACAGATGCTCAACAGATGTTCTCGATGTTGGAGAAATTTGCTTTAGTACATACATTT 2040  
 Db 1994 ACAGATGCTCAACAGATGTTCTCGATGTTGGAGAAATTTGCTTTAGTACATACATTT 2053  
 QY 2041 CTTGGCTTCACTGCTAGTACAGAGAAAGCTGTTGATAGTATGATCCAGAAAGGCG 2100  
 Db 2054 CTTGGCTTCACTGCTAGTACAGAGAAAGCTGTTGATAGTATGATCCAGAAAGGCG 2113  
 QY 2101 AAGAGCACTTTGAACACAGCTGGGTTTCTCTTCAAGAGCTCCCAATCCAGTAGATGAT 2160  
 Db 2114 AAGAGCACTTTGAACACAGCTGGGTTTCTCTTCAAGAGCTCCCAATCCAGTAGATGAT 2173  
 QY 2161 GAGTCATGGGAT 2172  
 Db 2174 GAGTCATGGGAT 2185  
 RESULT 2  
 AF262962 2411 bp mRNA linear PRI 30-MAY-2000  
 LOCUS AF262962 Homo sapiens VASA protein mRNA, complete cds.  
 DEFINITION AF262962.1 GI:8102020  
 VERSION AF262962.1 GI:8102020  
 KEYWORDS Homo sapiens.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 2411)  
 Rocha, D. and Affara, N.  
 Cloning and characterization of the human VASA gene  
 Unpublished  
 2 (bases 1 to 2411)  
 Rocha, D. and Affara, N.  
 Direct Submission  
 Submitted (02-MAY-2000) Department of Pathology, University of  
 Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK  
 JOURNAL TITLE Location/Qualifiers  
 FEATURES  
 source 1..2411  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 18..2192

/codon\_start=1  
 /product="Vasa protein"  
 /protein\_id="AF27205.1"  
 /db\_xref="GI:8102021"  
 /translation="MGDEWEAEINPHMSSVPIFEKDRYSGENGDNFNPRTPASSSM  
 DGPSPRDHPKSGFASGRNFGNRDAGECKRNTSTMGFGVGSFGNFGSRSRE  
 DGDSGSGFWRESNDCEMDPTNRNGFSRGGYRGDNSEAGPYRGRGSGFRGCRGF  
 GLGSPNNLDLPDECMQRTGGLFGSRPVLSGTGNDTSQSRSGSGSERGGYKINEV  
 ITSGKNSKMSSEAEAGESSDTQPKVTYIIPPPEDEDSIFAHYOTGINFKYDTILV  
 EVSGHAPPAIILFEEENILCOTLNINNAKAYTKLIPVOKVSIPIILAGDILMACOT  
 GSKRTAFILPIIAHMHDSITSRKELOPEPCTIVAPRELINOYILEARKSPET  
 CVAIVYIGGTQJGHSIROIVGCNLTICATPGRLMIDIGEKILGKQIKYLVDEAR  
 MDMGEPPEMKKLISCEPMPSKROTLMSATFPEBIOLEAEFLKSNLFAVVGAF  
 GCACRVVOQTVLQVGFQFSKREKLEILRNIGDRTFVFEYTKKADFIATFLCOEKIS  
 TTSIHGBREOREALGDRPFKCPVLVTSVAAGLDIENOVHINFLPSTIDEX  
 VHRIGTRGCGNTGRAISFPLDESMDHLAQPLVKVLTDAQODVPAMLEELAFSTYIPG  
 FSGSTRGNVNVASDTRKGSTLTITPAGESSSQANPVDDESWD"  
 BASE COUNT 771 a 408 c 571 g 661 t  
 ORIGIN  
 Query Match 99.6%; Score 2164; DB 9; Length 2411;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2167; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 ATGGGGGATGAAGATTGGGAAGCAGAATCAACCTCATATGCTTCTTATGTTCCATA 60  
 Db 18 ATGGGATGATGAAGATTGGGAAGCAGAATCAACCTCATATGCTTCTTATGTTCCATA 77  
 QY 61 TTGGAGAGATATGATATTTCTGAGAAAATGAGACAAATTTTAAACAGAGCTCCAGCTTCA 120  
 Db 78 TTGGAGAGATATGATATTTCTGAGAAAATGAGACAAATTTTAAACAGAGCTCCAGCTTCA 137  
 QY 121 TCATCAGAATATGATATGATGAGACCTTCGAAAGATATTTCAATGAAAGTGGATTGGC 180  
 Db 138 TCATCAGAATATGATATGATGAGACCTTCGAAAGATATTTCAATGAAAGTGGATTGGC 197  
 QY 181 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGTGAGTGAATAAAGCAGATAATATATCC 240  
 Db 198 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGTGAGTGAATAAAGCAGATAATATATCC 257  
 QY 241 ACAATGGGTGTTTGGAGTTGGAAAAGCTTTTGGAAAACAGAGCTTTTCAACACAGAG 300  
 Db 258 ACAATGGGTGTTTGGAGTTGGAAAAGCTTTTGGAAAACAGAGCTTTTCAACACAGAG 317  
 QY 301 TTTGAAGATGATATGATGATGAGCTGTTCTGAGAGAGCTGTGATAATGCTCGAAGATTAAT 360  
 Db 318 TTTGAAGATGATATGATGATGAGCTGTTCTGAGAGAGCTGTGATAATGCTCGAAGATTAAT 377  
 QY 361 CCAACACGAAACAGAGGTTTCCAAAGAGAGCGGCTATCGAGTGAATAATTTACAGAA 420  
 Db 378 CCAACACGAAACAGAGGTTTCCAAAGAGAGCGGCTATCGAGTGAATAATTTACAGAA 437  
 QY 421 GCTTCAGGGCCATACAGAGAAGGTGAAAGAGTAGTTCCGAGGTTGCCGTGAGAGATT 480  
 Db 438 GCTTCAGGGCCATACAGAGAAGGTGAAAGAGTAGTTCCGAGGTTGCCGTGAGAGATT 497  
 QY 481 GGTCTAGGAAGTCCAAATATGACTTAGACCCAGACGAATGTATGACGCGCACGTGTGGC 540  
 Db 498 GGTCTAGGAAGTCCAAATATGACTTAGACCCAGACGAATGTATGACGCGCACGTGTGGC 557  
 QY 541 CTTTGTGTTCTAGAAAGACAGATTAATAGTGGACAGATGATGATGATGATGATGATGATGAT 600  
 Db 558 CTTTGTGTTCTAGAAAGACAGATTAATAGTGGACAGATGATGATGATGATGATGATGATGAT 617  
 QY 601 AGAAGTGGCAGTGAAGTGAACAGAGTGTGTTACAAAGCTTTAAATGAAGAAATATACAA 660  
 Db 618 AGAAGTGGCAGTGAAGTGAACAGAGTGTGTTACAAAGCTTTAAATGAAGAAATATACAA 677  
 QY 661 GGCTCTGGAAGAAATTTTGGAGTCAAGACGAGAGAGAGAGAGAAATAGTATGATCTCAA 720  
 Db 678 GGCTCTGGAAGAAATTTTGGAGTCAAGACGAGAGAGAGAGAGAAATAGTATGATCTCAA 737  
 QY 721 GGACCAAAAGTGAATCATATCCCTCTCCACCTGAGATGAGAGCTCATCTTTTGA 780







```

Db      1861  TTTCTTCTTCAAGCTCCCAATCCAGATGATGATGATGATGAT 1907
|||||
RESULT 4
S75275 3030 bp mRNA linear ROD 15-MAY-1995
LOCUS   RVLG=vasa-like gene protein [rats, Wistar-Imamishi, testis, mRNA,
DEFINITION
3030 nt].
ACCESSION
S75275 GI:806463
VERSION
S75275.1
KEYWORDS
SOURCE
Rattus sp. testis Wistar-Imamishi.
Eukaryote. Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 3030)
AUTHORS Komiya,T. and Tanigawa,Y.
TITLE   Cloning of a gene of the DEAD box protein family which is
JOURNAL specifically expressed in germ cells in rats
MEDLINE Biochem. Biophys. Res. Commun. 207 (1), 405-410 (1995)
PUBMED  95160706
REMARK  GenBank staff at the National Library of Medicine created this
entry [NCBI g159371 from the original journal article.
This sequence comes from Fig. 1.
FEATURES
Source
1..3030
/organism="Rattus sp."
/db_xref="taxon:10118"
1..3030
/gene="RVLG"
/feature="putative DEAD box protein; This sequence comes from
Fig. 1; RVLG protein"
/codon_start=1
/product="Vasa-like gene protein"
/protein_id="AAB33364.1"
/db_xref="GI:806464"
/translation="MGDEDMEAELIKPHVSSYVVFPEKDKYSGANGDPTNRTSASS
EMEDGSGRDMRSGPSSGNLGNRDISSSKRETTSTGTFGRGKGFGRGLNKK
FERGDSGFMKSTNDCEDPTROTNRGRGKGYDPDNDSEASGPPRRGRDSEYDOD
GSORGGGLFGSRKPAASDSGSDPTFQSSGNARCAKYKLANEYVTVSSKSKWEABE
GESSDIQGPVTVIIPPEDEDESIPIAYOTGINPDKYDILIVSGHADAPALITE
EALNQCOTLNNNIAKAGYTKLIPVOKYSIPVLAGRDLMACQTSKTAFLPLIFLH
MRMDGITASRKEIAOEPECIIAPTRILINOIYLEARFSGTCVRAVIVYGTQFGH
SIRIOVOCNITLCATPGLMDIIGEKELIKOVKYLVDKADRMMDMGFGEMKLIS
CGMPSKROROTLLFSATPPEIOLRAGEPLKSNVLPVAVOVGAGACRDVQOSILOYG
PVFKRKUVELIRNIGDERPVPVETKKKAPFIATFLQDEKISTTSHGDEQERED
ALGDFRCGCPVLAVTSVAARGLDIENVQHINFLPSTIDIEYHRTGRKCNSTGR
AISEPDESNDHLAQPLVKVLSDAQODPAVLEIAPFSVAPSPFSNSTRGAVPASFD
TRKFKQGNITNTAGISSAQAPNPDEDSMD"
BASE COUNT 900 a 551 c 723 g 856 t
ORIGIN
Query Match 72.6%; Score 1577.6; DB 10; Length 3030;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 1862; Conservative 0; Mismatches 259; Indels 69; Gaps 6;

```

```

Qy      175  TTTGCCCTTGGGCGGAATTTTGGAAAAGAGATGCTGTGATGATATTAAGCAGATTAAT 234
Db      446  TTTTCCCTCGGAAGAAATTTAGGAAACAGAGATATTGGCGAGTCTGATTAAGAGAGACT 505
Qy      235  ACATCCAAATGGGTGTTTGGAGTTGGAAAGATTTTGGAAAACAGAGTTTTCAAAC 294
Db      506  ACATCTTCAACCCGTGCTTTGGAGAGGAAAGGTTTGGAAAACAGAGTTTTCAAAT 565
Qy      295  AGCAGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
Db      566  AACCAAGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 625
Qy      355  GATATTCGAACACGGAACAGAGGTTTCCAGAGAGCGGCTATCGAGATGGAATTAAT 414
Db      626  GATATCTAGACTGGAGCAGAGGTTTCCAGAGAGCGGCTATCCAGATGGAATGAT 685
Qy      415  TCAGAGCTTCAGGGCCATACAGAGAGTGGAAAGGTAGTTCCGAGGTTGCCGTGA 474
Db      686  TCGGAAGCTTCAGGCCCATTCAGAGAGTGGGAGAG----- 722
Qy      475  GGATTTGCTAGGAAGTCCAAATTAATGACTTGAACCCAGACGAATGTATCGAGCGACT 534
Db      723  -----ATAGTAATATGACCAAGATCAGGATCACAGCTGCT 760
Qy      535  GGTGGCTTTTGGTTCTAGAGACAGATATTAAGTGCACAGTAAATGATGATCTTCT 594
Db      761  GGTGGCTTTTGGTTCTAGAGAACAGAGCAAGTATTCAGGCAATGATGATGATCTTCT 820
Qy      595  CAAGCAGAGTGGCACTGGAAGTGAACGAGTGGTTTCAAGGTTTAAATGAAGAGTA 654
Db      821  CAGAGCAGA-----AGTGGAAATGCGCGAGGTGCTTCAAGGCTTAAATGAAGAGTA 874
Qy      655  ATACAGAGCTCTGGAAGATTTCTTGAAGTCAGAGCAGAGAGGAGAGAGAGATGAT 714
Db      875  GTAACAGGCTCTGGAAGATTTCTTGAAGTCAGAGCAGAGAGGAGAGAGAGATGAT 934
Qy      715  ACTCAAGAGCAAAAGTGAACCTACATACCCCTCTCCACCTGAGATGAGACTCCATC 774
Db      935  ATTCAAGTCCAAAGATGATATATACCCCTCTCCACAGAGATGAGACTCCATC 994
Qy      775  TTGCAATTAATCAGACAGCATTAATCTTGACAAATACGACTATTTTGTGGAATG 834
Db      995  TTGCAATTAATCAGACAGCATTAATCTTGACAAATACGACTATTTTGTGGAATG 1054
Qy      835  TCTGCAATATATGACACCAAGAACTTCTGACTTTGAAGAGTAATCTCTGACAGA 894
Db      1055  TCTGCAATATATGACACCAAGCAATTTGACTTTGAAGAGTAATCTCTGACAGACC 1114
Qy      895  CTGAATTAACAACTTGTAAAGCTGTTATATAGCTTACTCTCTGCAAAAATACAGT 954
Db      1115  CTGAATTAACAACTTGTAAAGCTGTTATATAGCTTACTCTCTGCAAAAATACAGT 1174
Qy      955  ATTCTTATCATCTTGAAGACGAGATTTGATGCTGTGCTCAACAGAGCTTGGGAG 1014
Db      1175  ATTCTTATGTTAGTGAAGAGATTTGATGCTGTGCTCAACAGAGCTTGGGAG 1234
Qy      1015  ACTGGGCTTTTCTTCAACAAATTTTGGCTCATATGATGATGATGATGATGATGATGAT 1074
Db      1235  ACTGGGCTTTTCTTCAACAAATTTTGGCTCATATGATGATGATGATGATGATGATGAT 1294
Qy      1075  CGTTTAAAGTGGAGAAACAGAGTATTTATGAGCAACCACTCGAATTTGGTCT 1134
Db      1295  CGTTTAAAGTGGAGAAACAGAGTATTTATGAGCAACCACTCGAATTTGATC 1354
Qy      1135  AACCAATTAATTTGGAAGCCAGAAAATTTTCTTTGGAGCTTGTGAAGCTGTTGT 1194
Db      1355  AACCAATTAATTTGGAAGCCAGAAAATTTTCTTTGGAGCTTGTGAAGCTGTTGT 1414
Qy      1195  ATATATGGGGAACCCAGCTGGGACATTTCAATTTGACAAATAGTACAGGCTGTAATTA 1254
Db      1415  ATATATGGGGAACCCAGCTGGGACATTTCAATTTGACAGATATGACAGGCTGTAATTA 1474

```





Query Match 66.4%; Score 1443; DB 10; Length 1930;  
 Best Local Similarity 86.4%; Pred. No. 0;  
 Matches 1671; Conservative 0; Mismatches 240; Indels 24; Gaps 6;

QY 253 TTGGAGTTGGAAGAGTTTGGAAACAGAGTTTTCACACGAGCTTGAAGATGT 312  
 DB 1 TTGGAGAGGAAAGGAGCTTTGGAAACAGAGTTTTCACACGAGCTTGAAGATGT 60

QY 313 GATAGCTGTGTTCTGGAGAGAGTCTAGTAATGACTCGCAAGTAATCCACACGAGAC 372  
 DB 61 GATAGCTGTGTTCTGGAGAGAGTCTAGTAATGACTCGCAAGTAATCCACACGAGAC 120

QY 373 AGAGGGTTTCCAAAGAGAGCGGCTATCGAGATGGAATAATTCAGAACTTCAGGGCCA 432  
 DB 121 AGAGGGTTTCCAAAGAGAGTGGCTGCGCAAGATGGAACCATTCAGAACTTCAGGGCCC 180

QY 433 TACAGAAAGAGTGAAGAGTGTTCGAGCTTCCGAGTTCGCGTGAAGATTTGGTCTAGGAAGT 492  
 DB 181 TTCAGAAAGAGGGAAGAGGAGCTTTCGAGGCTGCGTGAAGATTTGGTCTAGGAAGA 240

QY 493 CCAATATGACTTACACCCAGACGAATGTATGCAAGCGCACTGGTGCCTTTTGGTCT 552  
 DB 241 CCAATATGACTTACACCAAGATCAGGGGACACAGTGTGTGCTTCTGGTCTTA 300

QY 553 AGAAGACCAATTAATTAAGTGCACAGTAATGTGATATCTTCACAAACAGAAAGTGCAGT 612  
 DB 301 GGAAGAACCAACAGCAAGTGAATTCAGGCAATGTGACACTTACCACAAACAGAAAGTGCAGT 354

QY 613 GGAAGTGAAGAGTGTTCACAAAGTTTAAATGAAGAAATTAATACAGGCTCTGGAAG 672  
 DB 355 GGAAGTGTGAGAGTGTTCACAAAGTTTAAATGAAGAAATTAATACAGGCTCTGGAAG 414

QY 673 AATTTCTGGAAGTCAAGAGAGAGAGAAAGTGTATCTCAAGACCCAAAAGT 732  
 DB 415 AATTTCTGGAAGTCAAGAACTGAAGAGGTGAAGACGTATGATCAAGTCCAAAAGT 474

QY 733 ACCTACATACCCCTCTCTCCAGCTGAGATGAGAGCTCATCTTTGCATATTCAGACA 792  
 DB 475 ACATATATACCCCTCTCCAGCTGAGATGAGAGCTCATCTTTGCATATTCAGACA 534

QY 793 GGAATTAATCTTCACAAATACGACATATCTTGTGAAGTGTCTGACATGATGACCA 852  
 DB 535 GGAATTAATCTTCATTAATATGACACCATCTTGTGAAGTGTCTGACATGATGACCA 594

QY 853 CCAGCAATCTGACTTTTGAAGAGTAACTCTGTGACACATGAAATCAACATTCGT 912  
 DB 595 CCAGCAATTTTGACTTTTGAAGAGTAACTCTGTGACACATGAAATCAACATTCGT 654

QY 913 AAAGCTGTATATCTAAGCTTACTCTGTGCAAAAATACAGTATCTCTATCATATTCGA 972  
 DB 655 AAAGCTGTATATCTAAGCTTACTCTGTGCAAGATACAGTATCTCTATTCATTCGA 714

QY 973 GGAACGAGATTGATGCTTGTGCTCAAAACAGGCTCTGGAAAGCTGGGCTTTTCTCTA 1032  
 DB 715 GGAACGAGATTGATGCTTGTGCTCAAAACAGGCTCTGGAAAGCTGGGCTTTTCTCTT 774

QY 1033 CCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1092  
 DB 775 CCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834

QY 1093 GAAACGAGATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152  
 DB 835 GAAACGAGATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894

QY 1153 GCCAGAAAATTTTCTTTTGGAGCTTGTGAAGCTTGTATATATGAGGGAAACCCAG 1212  
 DB 895 GCCAGAAAATTTTCTTTTGGAGCTTGTGAAGCTTGTATATATGAGGGAAACCCAG 954

QY 1213 CTGGAGCATTCATTCGACAAATAGTACAGGCTGTAAATATATATATGATGATGATGATGATGATGAT 1272  
 DB 955 TTGGATGATTCATTCGACAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014

QY 1273 AGACTGATGATATCATATGACAAAGAAAGATTTGGTCTCAACAGATCAAAATACTTATGTT 1332

DB 1015 AGCCTATGACATCATAGCCAAAGAAAGATTGGGCTCCAAACAGTCAAGTCTTATGTT 1074

QY 1333 TTGGATGAGCTGATGCAATGTTGATGATGAGGTTTGGTCCAGAAATGAAGAAATTAAT 1392

DB 1075 TTGGATGAGCTGATGCAATGTTGATGATGAGGTTTGGTCCAGAAATGAAGAAATTAAT 1134

QY 1393 TCTTGGCCAGGAATGCAATCAAGAAAGAAACAGCCCAACCTTATGTTAGTGAACCTTTT 1452

DB 1135 TCTTGGCCAGGAATGCAATCAAGAAAGAAACAGCCCAACCTTATGTTAGTGAACCTTTT 1194

QY 1453 CCAGAGAAATTAAGAGTTGGCTGCAAGATTTTAAAGTCAATATCTGTTTGTCT 1512

DB 1195 CCAGAGAAATTAAGAGTTGGCTGCAAGATTTTAAAGTCAATATCTGTTTGTCT 1254

QY 1513 GTTGGCAAGTGGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572

DB 1255 GTTGGCAAGTGGGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314

QY 1573 TTCTCAAAAAGAGAAAGCTCGTTGAATTTCTGCG-AAACATAGGGGATGAAGAACTAT 1631

DB 1315 -TATCAAAAAGAGAAAGCTGTTGAGATTTCTACGAAACATAGGATGAAGAACTAT 1373

QY 1632 GGTCTTTGTTGAACCTAGAGAAAGCAAGATTTTACTGCAACTTTTCTTGTCAAGAAA 1691

DB 1374 GGTCTTTGTTGAACCAAGAAAGCAAGATTTTACTGCAACTTTTCTTGTCAAGAAA 1433

QY 1692 AATATCAACTACAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1751

DB 1434 AATATCAAGTACAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1493

QY 1752 AGATTTTGGCTTGAAGAGTCCAGATCTTGTGTTGACTTCAAGTACTCCAGAGGCT 1811

DB 1494 AGATTTTGGCTTGAAGAGTCCAGATCTTGTGTTGACTTCAAGTACTCCAGAGGCT 1553

QY 1812 GGAATTAAGAAATGTCGAACATGTTATCAATTTTGAATCTTCTTCTTACATGATGAATA 1871

DB 1554 TGAATTAAGAAATGTTCAACATGTTATCAATTTTGAATCTTCTTCTTACATGATGAATA 1613

QY 1872 TGTTCATCGAATTTGGGCGTACTGCTGTTGTGGGAATCTCGGAGAGCAATTTCCCTTTT 1931

DB 1614 TGTTCATCGAATTTGAGCGCACTGGCGCTGTGGAAATCTCGGAGAGCAATTTCTTTT 1673

QY 1932 TGAATCTTGAATGCGATTAACCATTTAGACAGGCTCTAGTAAAGTATTTGACAGATGCTCA 1991

DB 1674 TGAATCTGACTCTGATATATCATTTAGACAGGCTCTAGTAAAGTATTTGACAGATGCTCA 1733

QY 1992 ACAGAGTGTCTCGATGCTTGGAGAGAAATGCTTTAGTACATACAT--TCTGGCTT 2048

DB 1734 ACAGAGTGTCTCCGCAATGGCTAGAGAGATTTGCTTCAATACCTATGTGCTCCAGCTT 1793

QY 2049 CAGTGTAGTACAGA---GGAAGCTGTTTGCATCAGTGTATCCAGAAA----- 2096

DB 1794 CAGTGTAGTACAGAAGAGGGGGGCGTGTGTGATCTGTGACACAGAGGAATTAACA 1853

QY 2097 -GGGCAAGAGCACTTTGAACACAGCTGGGTTTCTTCTTACAGAGCTCCCAATCCAGTAG 2155

DB 1854 GGGCAAGAGCACTTGAATATCAGCGGGGATTTCTTCTTACAGAGCTCCCAATCCAGTTG 1913

QY 2156 ATGATGATCATGGG 2170

DB 1914 ATGATGATCATGGG 1928

RESULT 7  
 AB004836  
 LOCUS 1989 bp mRNA linear VRT 01-SEP-2000  
 DEFINITION Gallus gallus mRNA for Cvh, complete cds.  
 ACCESSION AB004836  
 VERSION AB004836.1 GI:967267  
 KEYWORDS Cvh.  
 SOURCE Gallus gallus  
 ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.

1 (sites)  
Tsunekawa, N., Naito, M., Sakai, Y., Nishida, T. and Noce, T.  
Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells  
Development 127 (12), 2741-2750 (2000)

2 (sites)  
Tsunekawa, N., Fujimoto, H., Nishida, T. and Noce, T.  
Isolation of a DEAD-family protein gene that encodes a chicken homolog of Drosophila vasa and its specific expression in germ cell lineage

Unpublished  
3 (bases 1 to 1989)  
Tsunekawa, N.  
Direct Submission  
Submitted (13-JUN-1997) Naoki Tsunekawa, Mitsubishi Kasei Institute of Life Sciences, Molecular Reproduction and Development; 11 Minamiooya, Machida, Tokyo 194, Japan  
(E-mail: tsunek@libra.18.m.kagaku.co.jp, Tel: +81-427-24-6246, Fax: +81-427-24-6314)

Location/Qualifiers  
1. 1989  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
1. 1989  
/product="CvH"  
/codon\_start=1  
/protein\_id="BAB12337.1"  
/db\_xref="GI:9967268"

BASE COUNT 597 a 385 c 493 g 514 t

ORIGIN

Query Match 34.6%; Score 751.2; DB 5; Length 1989;  
Best Local Similarity 71.1%; Pred. No. 8.6e-167;  
Matches 1009; Conservative 0; Mismatches 408; Indels 3; Gaps 1;

648 AGAAGTAATAACAGCCTCTGGAAGAATCTTGGAGTCAGAGACGAGAGAGAAAG 707  
Db 492 AGGAGGAAG 551  
Qy 708 TAGGATAGTCAAG 767  
Db 552 TGATCTGGAAG 611  
Qy 768 CTCATCTTTGACATTAATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827  
Db 612 GTTCATCTTTGACATTAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 671  
Qy 828 GGAAGTGTCTGAGACATGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 887  
Db 672 TGAGATGTCAAG 731  
Qy 888 TCAAGACATGAATTAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 947  
Db 732 TCAGACTTTAAG 791  
Qy 948 ATACAGATTCCTTCACTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007

Db 792 GCACAGATTCCTGTTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 851  
Qy 1008 TGGGAAG 1067  
Db 852 AGGAAAAAG 911  
Qy 1068 TGGCAGTCTTTTAAAG 1127  
Db 912 TGCAGC--TTCCAAAG 968  
Qy 1128 ATTGTCACAGAGATTTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187  
Db 969 ACTGATTAATTCAGATTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1028  
Qy 1188 TGTGTTATATATAGGGAG 1247  
Db 1029 TGTGTTATATATAGGGAG 1088  
Qy 1248 TAT 1307  
Db 1089 TAT 1148  
Qy 1308 TCTCAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1367  
Db 1149 TTTGTTGAG 1208  
Qy 1368 TGGTCCAG 1427  
Db 1209 TGGATTAAT 1268  
Qy 1428 AACCTTAT 1487  
Db 1269 AACTTAT 1328  
Qy 1488 AAGTCAAT 1547  
Db 1329 GAAAG 1388  
Qy 1548 GCAG 1607  
Db 1389 GCAAAAT 1448  
Qy 1608 AAACATAG 1667  
Db 1449 AAGCAG 1508  
Qy 1668 TGCACATTTCTTTGTCAG 1727  
Db 1509 TGCAGCCTTTCTTTGTCAG 1568  
Qy 1728 GAG 1787  
Db 1569 GAG 1628  
Qy 1788 TACTTCAATAGTCCAG 1847  
Db 1629 AACTTGGATGATCAAG 1688  
Qy 1848 TCTTCTCTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1907  
Db 1689 TCTTCTCTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1748  
Qy 1908 TACTTGGAG 1967  
Db 1749 TACTTGGAG 1808  
Qy 1968 AGTAAAGATTTGACAG 2027  
Db 1809 ACTTAAAGTCTTTCCAG 1868  
Qy 2028 TAGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2067  
Db 1869 AAGAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1908





Qy	1767	AAAGTCCCAATTCCTGTGCTACTACTAGTAGCTCCAGAGGCGCTGATATTGAAAAATGT	1826
Db	1790	AAAGGTAAGTATTATGTCTGCACACAGCTTGCTCCAGAGGCTTGATATGAAAAATGT	1849
Qy	1827	GCAACATGTTATTCATTTTGTATCTCTTCCTTACACATGATGATAATGTTTCAGAAATTGG	1886
Db	1850	TCAACAGCTGATTAATTTATGACGTTCCCTAAGGAAAGTTGATGATGATCGTCCATTAGAAATTGG	1909
Qy	1887	GCGTAAGTGTGCTGTGGGAATACTGCGAGAGCAATTTCCCTTTTGTATCTTGAATCGGA	1946
Db	1910	TCTGATCCGTCGTGTGGTAACACCGGAAAGGCAACATCAATTTTCAATGTTCCAGATGA	1969
Qy	1947	TAACCATTTACACAGCCTCTAGTAAAGTATTGACAGATGCTCAACAGATGTTTCCCTGC	2006
Db	1970	CCAGTGTATGCTGTGCCCTTTGGGAAATTTCTTAACGATGCTATCAAGAAAGTCCCTGC	2029
Qy	2007	ATGCTGGAGAAATTTGCCCTTTAG	2030
Db	2030	TTGCTTAAAGAAATTTGCCCTTTGG	2053

RESULT 9			
AB005147			
LOCUS	AB005147	2665 bp	linear
DEFINITION	Danio rerio mRNA for vas, complete cds.		VRT 01-OCT-1997

VERSION	AB005147.1	GI:2463518
KEYWORDS	vas.	
SOURCE	Danio rerio cDNA to mR	

Euharyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE	AUTHORS	TITLE
1	(sites), Yoon, C., Kawakami, K. and Hopkins, N.	Zenaristish vaa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells

JOURNAL Development 124 (16), 3157-3165 (1997)  
MEDLINE 97417583  
REFERENCE 2 (bases 1 to 2865)  
AUTHORS Kawakami, K.  
TITLE Direct Submission  
JOURNAL Submitted (25-JUN-1997) Koichi Kawakami, The Institute of Medical  
Science, The University of Tokyo, Department of Tumor Biology;  
Shiroanadai 4-6-1, Minato-ku, Tokyo 108, Japan  
(E-mail:kawakami@ims.u-tokyo.ac.jp, Tel:81-3-5449-5308)

FEATURES	Location/Qualifiers
source	1. .2865

```

/organism="Danio rerio"
/db_xref="taxon:7955"
71..2221
CDS

```

```
/codon_start=1
/product="vas"
/protein_id="BAA22535.1"
/db_xref="GI:2463519"
```

/translation="MDDWEEDQS  
EGTEGSSWKMTGDSFRGRGGGG

GFRGQFRRSGEDENRNDNDGMMKGBESRRGRRGQFGSGFPGGRRGDRNEDJOTRRGRCG  
 RENNNGNDEDEGGRRGRRGFRGQFRGQDEGGRRGFRGGRGRNIEEYVFSKTTTAA  
 DLDOEGSANGKRVVVVPPPEEBEESIFSHYVINGNFDYDIDLVDVSSNGPKAL  
 MTPBEGJCDLSLKNVSKSGVYVPPVPOKHQIPIISAGRIDNCAQNGSGKTAFLPIL  
 ILOBPMTDGVASKEPILQEBEALIVAPETRLINOJILEAKFAGVNCVRVVVYGGI  
 NTGTYIRVLKGCNVLCATPGRHLHDLGRGKTSKRYILVBDAAVNDGSDVEPBNRR  
 KLVASPGMRSEKEBRTLMFSATYIPEDIOKMAADPLKVDYIPLAVGANGGASDVEQVY  
 VQVOYRSKDQJLLELRATGNEKRWVPEYTRASDPAFLTCQKISHTSIHGRBER  
 EREKALSDFRJGHCPEVATVSNVAGLDIEOVHVNEMDFMSIDEVYHARGRGRCG  
 TNDGAVSFNFNPSEDTPIARSLVKVYLSGNOQVVPMLBEVASFAGTGTGFNRGCVFVAS  
 TDSRGSGSKFDEDEPPQTSAPASAAAAADPEWE

BASE COUNT	783 a	525 c	791 g	766 c
ORIGIN				

Query Match 31.6%; Score 686.8; DB 5; Length 2865;

Best Local Similarity 68.6%; Pred. No. 1.4e-151;  
Matches 959; Conservative 0; Mismatches 436; Indels 3; Gaps 1

Qy	702	AGAAAGTAGATACTCAAGGACCAAAAGTACCAATACCCCTCTCCACATGAGGA	761
Db	739	AGAAAGAGTAAATGACAGACCCAAAGTGTATTATGTGCGCCGCTCTCCAGAGAA	798
Qy	762	TGAGACTCCATCTTTGACATTAACAGACAGCATTAACCTTCGAAATACGACATAT	821
Db	799	GGAGAGCTCGATATTTCTCCATTATGCAACAGGCAATTAATTTTGCAAAATATGATGATAT	858
Qy	822	TCTTGGAAGTGTCTGAGCATGATGCAACCAAGCAATTCGATCTTTTGAAGAAGCTTAA	881
Db	859	TCTTGGAAGTGTGAGTGCAGCAATCTTCCAAAGCAATTAATGACTTTTGAGAAAGCAGG	918
Qy	882	TCTCTGTACAGACTGTAATTAACAACTTGTAAAGCTGTATATACAGCTTACTCTGT	941
Db	919	ACTTTGTGACTCAGTACGAGCAAAAATGATCAAAAGCTGGATATGGAAGCCTACACTGT	978
Qy	942	GCAAAAATACAGTATTTCTATCATACTTCGAGACGAGATTTGATGCTGTGTCTCAAC	1007
Db	979	TCAGAAATGTGCAATTTCCCATCATTTTCAGCTGGACGGGATCTAAAGGCTTGTGCTCAGAC	1038
Qy	1002	AGGCTCTGGAAGAAGCTGGCGCTTTTCTCTCAATTTTGGCTCATATGATGATCAGATG	1065
Db	1039	TGATACAGAAAAAGCGCGCTTCTCTGCTCTATCTACAGCGCTTATATGACTGATG	1099
Qy	1062	AATACTGCCAGTCGTTTTTAAAGATTGCAGGAACAGAGTGTATTAATGTAGACCAAC	1122
Db	1099	TGTGGACAGCCAGCAAGTTCAGTGAAGATACAGAGAGCTGAGGCCATTAATGTGTGCTCCAC	1158
Qy	1122	TCGGAATTTGTCACACGAATTTATTTGGAAGCCGAATTTTCTTTGGGACTTGTGT	1188
Db	1159	CAGAAACTTAATCAATCAGATCTATCTGGAAGCCAGGAATTTGCTATACGGGACCTGTGT	1211
Qy	1182	AAGAGCTGTTTATATATATGGGGGAACCCAGCTGGGACATTCATTTGACAAATAGTACA	1243
Db	1219	ACGACCTGTGTGTTATATGAGGATATAAATACATGATATACATCCAGAGGTGTTAA	1276
Qy	1242	AGGCTGTATAATATATGTGCTACTCTGGAAGACTGATGATATCATAGCAAAAGAAA	1307
Db	1279	GGGCTGCATGTCTGTGTGCTACTCCCTGGAAAGATTGATGACCCATATGGTGTCTGGAAA	1338
Qy	1302	GATTGCTCAACAGATCAAAATCTTATGTTTGGATGAAGTGATGCGATTTGSGATAT	1366
Db	1339	GATTGGCTTGATTAAGTGCCTATCTAGTTCTGATATAGCAGACAGATGCTGACAT	1399
Qy	1362	GGGTTTTGTGCAGAAATGAAGAAATTAAATTTCTTGCCAGAGATGCAATCAAGAAACA	1422
Db	1399	GGGCTTTGAGCAGAAATGCGCAAACTGCTGAGCGCTCTCGTATATGCTTCAAAAGAGGA	1455
Qy	1422	GGCGCAAAACCTTATATGTTCAGTGCACATTTTCCAGAGAAATTCAAAGTTGGCTGCAGA	1488
Db	1459	ACGACAAACCTTCATGTTCAGTGCACCACTCCGTGAATATTCAAAGAAATGGCTGCTGA	1511
Qy	1482	GTTTATAAGTCAATTAATCTGTTTGTGTGCTGTGGAACAAGTGGGTGAGCATGTATGAGA	1544
Db	1519	CTTCTTAAAGTGAATATATTTTCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1576
Qy	1542	TGTTTACAGCAGCCGTTCTCAAGTTGGCCAGTTCCAAAAAGAGAAAAGCTCGTGAAT	1607
Db	1579	TGTGAGCAAAACCGTTGTTCAAGTGGACCAATACCAAAAGAGGAGCCAGCTGCTGAAAT	1633
Qy	1602	TCTGCGAAACATAGGGGATGAAGAACTATGTCTTTTGAATAATGAAGAAAAGCAGA	1666
Db	1639	GCTCAGACCAACAGGTAAATGAGCCGACAAATGTTTTTGTGAACCAAAAGAAAGTGCCTGA	1699
Qy	1662	TTTTACTCAACTTTTCTTTGTCAAGAAAAATATCAACTCAAGTATCCATGTGATCG	1722
Db	1699	TTTCAATAGCAATTTTCTGTCTCAAGAGAAATCTCAACCAACAGCATCCATGTGATG	1755
Qy	1722	GGAAACAGAGAGCGGGAGCAAGCTCTTGAAGATTTTGTGCTTTGAAAGTGCACGTTCT	1788



Db		1759	GGAACAGGAGAAGGAGAAAGCTTCAGTGGATTTTGCGCTTGCCACTGTCCGGTTCT	1818
Oy		1782	TGTTGCTAATTCACTAGTAGTGTGCCAGAGGGCTGGATTTGAAAATGTGCACAATGTTATCAA	1841
Db		1819	GGTGGCCACATCTGTAGCTGTCTAGAGGCGCTGATTAATGACAAGTCCAGCATGTTGTGAA	1878
Oy		1842	TTTTGATCTTCTTCTACCATTTGATGATATGTTTCATGGAATTTGGCGCTACTGGTCTGT	1901
Db		1879	TTTGAATGCCCCACAGCAGCATGTAGATATGTCCATGTCAATGAGAACACTGACGCTG	1938
Oy		1902	TGGGAATTAAGTGGCAGAGCAATTTCTCTTTTGTGATCTTGAATCGGATAACATTAGCACAA	1961
Db		1939	TGGGAACACCGGTGCGCGCGTGTCTTTTAAACCAGAGTGTACATCTCACATAGCTCG	1998
Oy		1962	GCTCTAGTAAAAAGTATTGACAGATGCTCAACAGAGATTTCTCTGATGTTGGAAGAAAT	2021
Db		1999	CTCCCTGGTCAAGAGTCTTTTCAGGGGCCCAACAAGTATGTTCCAAAATGGCTGAGGAAGT	2058
Oy		2022	TGCGTTTAAAGTACATATCAATTCCTGTGGTTAGTGTATGTAACAAGAGAAACGTTTGGATC	2081
Db		2059	TGCGTTCAAGTCTCA--TGSCACAACAGGGTTTCAACCCACAGTGGAAAGTGTGTTGCATC	2115
Oy		2082	AGTGATACCAAGAAAGG	2099
Db		2116	TACAGACTCGCGGAAGG	2133
RESULT	10			
LOCUS	DRY12007		2492 bp	mRNA linear VRT 09-OCT-1997
DEFINITION	D.rerio vlg (vasa like gene) mRNA for putative RNA helicase.			
ACCESSION	Y12007			
VERSION	Y12007.1	GI:2558534		
KEYWORDS	RNA helicase; vasa gene; vlg gene. zebrafish.			
SOURCE	Danio rerio			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.			
REFERENCE	1 Olsen, L.C., Aasland, R. and Fjose, A. (bases 1 to 2492)			
AUTHORS	Olsen, L.C., Aasland, R. and Fjose, A.			
TITLE	A vasa-like gene in zebrafish identifies putative primordial germ cells			
JOURNAL	Mech. Dev. 66 (1-2), 95-105 (1997)			
MEDLINE	98025484			
PUBMED	9376327			
REFERENCE	2 (bases 1 to 2492)			
AUTHORS	Olsen, L.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (24-MAR-1997) L.C. Olsen, University of Bergen, Department of Molecular Biology, Thormoehlenstgt 55, N-5020 Bergen, NORWAY			
FEATURES				
source	Location/Qualifiers			
	1..2492			
gene	/organism="Danio rerio"			
	/db_xref="taxon:7955"			
CDS	/dev_stage="embryonic, late blastula"			
	36..2138			
	/gene="vlg"			
	36..2138			
	/gene="vlg"			
	/note="vasa-like gene; putative"			
	/codon_start=1			
	/evidence=not experimental			
	/product="RNA_helicase [DEAD box]"			
	/protein_id="CAA72735.1"			
	/db_xref="GI:2558535"			
	/db_xref="SPTRMBL:O42378"			
	/translation="MDWDEDQSPVSCSGFGAGANDKNSSEGTGSSWMVTGDSFDF GRGGGSRGGRGGSFKEIDENSGDGNNGGSRGRGGRGGRGGRSGSDPBNDE NCNDWGCGSSRGRGGRGGRGGRGGRGDGDNEDTGRGGRGRNNENGNDGCGRGRR RGGRGFRGGRFDGGDESGKGFGRGGRGVNRNEVFVSVTTADKUDOGSENAGFKVA YVPPEEESIFSHVATGINFPDYDDILIVDGSGSNPKAIMTEFEAGLCDLSIKNV			

protein_bind									
2255..2277									
/note="similar to nanos response element"									
/bound moiety="a putative nanos-dependent factor"									
/evidence=not experimental									
BASE COUNT	703 a	453 c	703 g	633 t					
ORIGIN									
Query Match	31.6%;	Score 686.4;	DB 5;	Length 2492;					
Best Local Similarity	69.6%;	Pred. No. 1.8e-151;							
Matches 927;	Conservative 0;	Mismatches 405;	Indels 0;	Gaps 0;					
Qy	702	AGAAAGTAGTGATAC	TCAAGGACCAAAAGTAC	TACATACCCCTCTCCACCTGAGGA	761				
Db	656	AGAAAGAAAGTAAAT	CAGAGCCCAAGGTTGTTA	TGTCGCCGCTCTCCACGAGGA	715				
Qy	762	TGAGACTCCATCTT	TGCACATTA	TCAAGAGGATAACTTGACAAATACGACATAT	821				
Db	716	GGAGAGCTCGATAT	TCTCCATTA	TGCAACGAGATTAATTTGACAAATATGATAT	775				
Qy	822	TCTTGGAAGTGTCT	GGACATGATGACCA	CCAGCAATTCGACTTTTGAAGAGCTAA	881				
Db	776	TCTTGAGACGTGAT	GCGACGATCCTCCAAAAGCAATTA	TGACTTTTGAGGAAGCAGG	835				
Qy	882	TCTCTGTACAGACT	GTAATAACAACATTTG	TAAAGCTGTTATATACTAAGCTTACTCCTGT	941				
Db	836	ACTTTGTGACTCA	CTGAGCAAAAATGTATCAA	AGTCTGATATGTGAAGCTTACACCTGT	895				
Qy	942	GCAAAATACAGTAT	TCTCTATCATCTTGACAGACGAGATTTG	TATGCTTGCTGTCTCAAC	1001				
Db	896	TCAGAAACATGGCA	ATTCCTCATCTTCACTGACGAGGATCTA	TATGCTTGCTGTCTCAAC	955				
Qy	1002	AGGCTGTGGGAAG	ACTCGGCTTTTCTCTCAACAATTTTG	GCTCATATGATGATGATGATG	1061				
Db	956	TGGAATCAGGAAAA	CGCGCGCTTCTCTGCTCCATCTCA	GCGCTTTATGATGATGATG	1015				
Qy	1062	AATAACTGCCAGT	GTTTAAAGATTGCAGGAACAAGATG	ATTAATTGATAGCACCAAC	1121				
Db	1016	TGTGGCAGCCGCA	AGTTCAAGTAGATACAGGACCTG	AGGCCATTAATCGTGCTCCAC	1075				
Qy	1122	TCGAGAATTGAT	CAACGATTTATTTGGAAGCCAGAAAATTTT	CTTTGGGACTTGTGT	1181				
Db	1076	CAGAGAACTTAT	CATCAGATCTATCTGGAGGCCAGGAAGTTT	CATACGGGACTGTGT	1135				
Qy	1182	AAGAGCTGTTG	TATATATGGGGGACCCAGCTGGGACATTC	CAATTCGACAAATATGTACA	1241				
Db	1136	ACGACCTGTTGT	GTTTATGAGGATTAATAACCGGATATCTA	TCCGAGAGGCTTTAA	1199				
Qy	1242	AGGCTGTAATAT	TATATGCTACTCTCTGGAGAGCTGATG	ATATCAATAGGCAAAAGAAA	1301				
Db	1196	GGGCTGCAATG	ATATGATGCTACTCTGGAGAGTTGATG	ATGATGACCTCAATTTGTCGGAAA	1255				
Qy	1302	GATTGCTCTCAAC	AGATCAAATCTTAATTTTGGATGGAAGCTG	ATGATGCGATGTTGGATAT	1361				
Db	1256	GATTGGCTGAGT	AAAGTCGCTATCTTAAGTTCTGATG	AAACAACAAGAAATGCTGCAT	1315				
Qy	1362	GGGTTTTGTC	CAAAATGAAAGATTAATTTCTTGCC	CGAGAAATGCGATCAAAAGGAACA	1421				
Db	1316	GGGCTTTGAG	CCAAATAATGCGCAACCTG	TGGCGCTTCTGATATGCTTTAAAGAGAA	1375				
Qy	1422	GGGCAAAACCTTA	TATGTTCAAGTGCACATTTTCCAGAGAAATTC	CAAAGGTTGGCTGCAGA	1481				
Db	1376	ACGACAAACCTTA	TATGTTCAAGTGCACCTACCTG	AAAGATTTCAAAGAAATGGCTGCAGA	1435				
Qy	1482	GTTTTTAAGTCAA	ATTAATTA	CTGTTTGTCTGTTGACAGTGGGTGAGCATGTACAGA	1541				



Oy		1431	CCTTATGTCAGTCACTTTTCCAGAGAATAATCAAAGGTGGCTGCAGAGTTTTTAA	1490
Db		1163	TCTGATGTCAGTGCTACTTTCCCAGGAACATCCAGGCGTTGCAGCTGTACTTTAA	1222
Oy		1491	GTCAAATTATCTGTTTGTGCTGTGACAAGTGGGTGAGACATGTAGATGTTACACA	1558
Db		1223	GACTGACTATTGTTCTTGCGCCGTGGGATATTGTGGGTGAGCGTGCAGATGTGAGACA	1282
Oy		1551	GACCGTCTCCAAGTTGGCCAGTTCTCAAAAAGAAAAAGCTCGTTGAATTTCGCCAA	1610
Db		1283	GACATTGTGTGAGGTCACCAAGTTCCGAAAAGGGAGACGCTAATAGCCTTCGAAATC	1342
Oy		1611	CATAGGGGATGAAGAATACTATGTCCTTTGTTGAAAATTAAGAAAAAGACATTTTACTGC	1670
Db		1343	AACAGGATCGAGCGCACCATGAGTGTGTTGTAGAGCCAAAGAGACAGGCTGATTTTATTCG	1402
Oy		1671	AACCTTTCTTTGTCAAGAAAAAATAATCAACTACATCAAGTATCCATGGTGAATCCGGAAACAGAG	1730
Db		1403	CACAATCTGTGTGTCAGAGGAAGTTTCCAACTACAGACATTTACGCTGACCGTGAAGCACTG	1462
Oy		1731	AGAGCGGAGACAAGCTCTTGAGATTTTGCCTTTGGAAAAGTCCCAAGTTCTTGTGTAC	1790
Db		1463	GCACCGAGAGACAGGCCCTTGGAGACTTCGCGCTCGGCAAAATGTTCTGCTCGTAGGCACAC	1522
Oy		1791	TTCAAGTGCCTCCACAGAGGCTGATATTTGAATAAATGTGCACATGTAATCAATTTTGACT	1852
Db		1523	CTGTGTAGTGCCCGTGTGACTGATATTCAGATGTGTGAGTTGCTGTGGAACTTTGACCT	1582
Oy		1851	TCCTTTACATTTGATGAATATGTTTCATCAATGAGCGCTACGCTGTTGGGAATAC	1910
Db		1583	CCCCAAACATATCGACGAATATGTCATCTGATTTGTGTGAACCCGCGCTGTGGCAACAC	1642
Oy		1911	TGCGAGAGCAATTTCTTTTGTGATCTTGAATCGGATAACCATTTTACACACCCTTAGT	1970
Db		1643	CGGAGGGGCACTGCTCTTCTATGATCCAGAGCTGACAGTAGTTGCTCATCATCTGCT	1702
Oy		1971	AAAGTATGACAGATGCTCAACAGAGATGTTCTGTCAGTGTGGAAAGAAATTTGCTTTAG	2030
Db		1703	CACATCTCTGTCAAAGGCCACAGAGAGTTCCCTCGTGTGTAGAGAGTTTGGTTTCAAG	1762
<hr/>				
RESULT_12				
	LOCUS	ABO32467		
	DEFINITION	Oreochromis niloticus mRNA for vasa, complete cds.		
	VERSION	ABO32467		
	KEYWORDS	ABO32467.1 GI:11990881		
	SOURCE	Vasa; vas.		
	ORGANISM	Oreochromis niloticus cDNA to mRNA.		
		Oreochromis niloticus		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
		Acanthomorphi; Acanthopterygii; Percormorpha; Perciformes;		
		Labiroidel; Cichlidae; Oreochromis.		
	REFERENCE	1 (sites)		
	AUTHORS	Kobayashi,T., Kajiyura-Kobayashi,H. and Nagahama,Y.		
	TITLE	Differential expression of vasa homologue gene in the germ cells during oogenesis and spermatogenesis in a teleost fish, tilapia,		
		Oreochromis niloticus		
	JOURNAL	Mech. Dev. 99 (1-2), 139-142 (2000)		
	MEDLINE	20544961		
	REFERENCE	2 (bases 1 to 2130)		
	AUTHORS	Kobayashi,T.		
	TITLE	Direct Submission		
	JOURNML	Submitted (16-SEP-1999) Tohru Kobayashi, National Institute for Basic Biology, Laboratory of Reproductive Biology, 38 Nishigunaka, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan		
		(E-mail:tohru_kemibb.ac.jp, Tel:81-564-55-7554, Fax:81-564-55-7556)		
	FEATURES	location/Qualifiers		
	source	1..2130		
		/organism="Oreochromis niloticus"		
		/db_xref="taxon:8128"		
	gene	1..2130		

CDS
48. .1985
/gene="vab"
/gene="vab"
/codon_start=1
/product="vasa"
/protein_id="BABI9807.1"
/db_xref="GI:11990882"
/translation="MDMEERQETRTATSCALTSIDTSTSGCTGDPWMNSDHGEFOGRRGG RKGGRGFKNVSFGSAGEDGANDGDQDNNTAERSFSRRRGRRGGRFGBTHSDPSF ANBEDGNKXNVSRRGRGGRGGQGQGGYRGDESLRLPKARYPENADASDR PRVLYVPPEPSEDESIFSHYESGVNPDKDELIVVDGSNPLPAINTPENAADSOR KRNVKSQYKPTPPVOHGHPISIASGDLMDVTSGSKTAAPLPILOOLMAAGAA SCSEMERPDALITVRRELINOIYLEARFAVGTCRPVVYGVSRVLTPDILGLG CNVICTPGRIPLDMIRGKXGIGLTRVIRLIVIDEADRMLDMGFEPMRLVSPGMSRE NRLOTMFSAFTFPEDIORLADFLKTDYLFLAVGIVGACSDVEOTFOVTVKPARBDL IDLKSTGSETRTWPFETRKOADFIALICOEKPPTSIGHDRPOMORMDLGPFRGS KCSLVATSVSGARGLDIPDVOPVNFDLPNNIDIEYVHRISTRGCNGTGAVALSYDFG ADSELARSILTITSKAOEVPMSLEEFASFHSISPSSGNPPRKHFASIDLRKPGCGFF QDSYSQLAOGTAADDEBEW"
BASE COUNT      560 a    458 c    624 g    488 t
ORIGIN
Query Match                  28.5%; Score 619.4; DB 5; Length 2130;
Best Local Similarity       66.8%; Pred. No. 1.2e-115;
-Matches 882; Conservative 0; Mismatches 438; Indels 0; Gaps 0;
Db      711 TGATACTCAAGAACCAAAAGTGACTCATATACCCTCCTCCACCTGTAGATGAGAATC 770
515 TGGAGTAGATAGACCAGGGTTACTCTAGTCCCCCACCACTCTGAATGAGAGTCC 574
Qy      771 CATCTTGCACATTATCAGACAGGCATTAACCTTGACAATAAGCACTATTCTGTGGA 830
Db      771 CATTTTCCCACATAGTACGCCGTGTCCAATTGACAAAGTACATAGATCTTTGTGGA 634
Qy      831 AGTGTGGAATGATGACCAACCAAGCAATCTGACTTTGGAAGATATCTCTGTCA 890
Db      831 TGTCAGTGGAACCAACCCACCACTATATATACATTTGATBAAGCAGCGCTGTGGA 694
Qy      891 GACACTGAATACCAACTTGCTTAAAGCTGGTTATATCAAGCTTACTCCTGTGCAAAATA 950
Db      891 GTCCCTAAAGAAAGGTCAGCAAGTGTGTTATGTBAAGCCACCCCTGTGCAAGAAACA 754
Qy      951 CAGTATTCCTATCATCTTGACAGACAGATTTGATGGCTTGTGCTCAACAGGGTCTGG 1010
Db      951 CGGCATCCCATCATCTTCTGCTGGCAGAGATCTCATGGCCGCGCCAGACATGGGCTTGG 814
Qy      1011 GAATACGTGGCGCTTTTCTCTTCCATCAATTTTGGCTCATATGATGATGATGAATTAATCG 1070
Db      815 AAAAAGGGCTGCTTCTGCTCCCAATCTTCAACACACTATGTGCAATGTGTGGCGGCG 874
Qy      1071 CAGTCGTTTTAAAGATGTGAGCAACACAGAGTGTATTGTGTAGACCAACATCCAGAAAT 1130
Db      875 CAGTTGCTTCAGTBAAGATGACAGAGCCTGATGCCATATGTGTGACCAACATAGGAGCT 934
Qy      1131 GGTCACACCAATTTATTTTGGAAGCAGAAAAATTTCTTTTGGGACTTGTGTAGAGCTGT 1190
Db      935 GATCAACCAAGATTACTCGTAGGGCCAGAAAGTTTGCTATATGGACATGTGTGCACTGT 994
Qy      1191 TGTATTATATGGGGGAACCCAGCTGGGACATTTCAATTGACCAATTAGTACAAGCTGTAA 1250
Db      995 TGTATGTATATGTGGAGTCAAGCACTGTGACACCAAATATAGAGACCTCTTGAAGGGCTGTAA 1054
Qy      1251 TATATTATGTGCTACTCTGTGAAGACTGATGATATCATAGGCAAGAAAAGATTGGCTCT 1310
Db      1055 TGTGTTGTGTGGAACACCAAGACGACTGTGTGACATATGTGGAAGAGAAAGTTGGCTT 1114
Qy      1311 CAACAGATCAAAATACTTATGTTTGGATGGAAGCTGATGCGATGTTGGAATATGGGTTTTGG 1370
Db      1115 GACTAAAGTGCGTTACTTGTGTGCGATGAGAGCTGACCGCATGTTGGATATATGGGGTTTGA 1174
Qy      1371 TTCGAAATGAAGAAGTTAATTTCTTGGCCAGGAATGCCATCAAGGAACAGCGCCAAAC 1430
Db      1175 GCCGTGACATGCGCGCGCTGTGGGCTTCTCTGGAGTGCATCAAAAGAACCGCTGCAGAC 1234



Db		1181	GACATCGGCTTTGAGCCTGCATCGCGCAAGCTGGTGGGTGCCCCAGGACATCACGATTAA	1240
Qy		1417	GAACGCGCCCAAACCCCTTAATGTTCAAGTGCACTTTCCAGAGGAAATTCAAAGTTGGCT	1476
Db		1241	GAGGATCGCCAGACCCTTAATGTTCAAGCGCTACCTAACCCCGAGCATCTCCAATAAATCGACT	1300
Qy		1477	GCAGAAGTTTTAAAGTCAAAATTAATCTGTTGTGCTGTGTGGACAAGTGGGGGAGCAATG	1536
Db		1301	GCTAATCTTCTGAAGAAAGACTTAATCTGTTCTTGTGTGTGGTGTGGTGGGAGCCTGC	1360
Qy		1537	AGAGATGTTCCAGCAGACCGTTCTCCAAATTGGCCAGTTCCTAAAAGAGAAAAGCTCGTT	1596
Db		1361	AGTATGTGGAGCAGAGTGTGTGTTCAGTAGCAAGTTCCTCCAAAAGAGCAACTACTG	1420
Qy		1597	GAAATTTTCGGAACATAGGGGATGAAGAACTATGTCCTTTGTGTAACCTAGAAAAAA	1656
Db		1421	GAGGTCTCTGAAGACTACAGGGTCTGAACGACACATAGGCTTTGTGGAAAACCAAGGCGAG	1480
Qy		1657	GCAATTTTACTGTGAACCTTTTCTTTGTGCAGAAAAAATATCACTACAGTATCCATGCT	1716
Db		1481	GCTAATCTTATAGCAACGTTCTCTGTGTGGAGAAGGTTAATCTACTAGTATTCACGCT	1540
Qy		1717	GATCGGGAACAGAGAGACGGGAGCAAGCTTTTGAGATTTTCCCTTTGGAAAATGCCCA	1776
Db		1541	GACCGTAGCAGAGAGGAGCGTGAACAGCGCTCCGCGACATTCGCTCCGGAAGGTGTCT	1600
Qy		1777	GTTCTTGTGCTACTTCAATAGTACTGCCAGAGGGCTGATATATGAAATGTGCCACATGT	1836
Db		1601	GTCTCTGGTGGCACCCTGTGTGTGCCCGCGAGCATGGAATCAAGATGTCCACACATA	1660
Qy		1837	ATCAATTTTGAATCTTCTCTTCAATGATGATGAATATGTTCATGCAATTTGGGCGTACTGCT	1896
Db		1661	GTCAACTTTGACCTCCGCCCAACAATATGATAGTACGCCACCCATCTGGGAGAAACAGT	1720
Qy		1897	CGTGTGGGAATATCGCAGAGCAATTTCCCTTTTGTATCTTGAATCGATTAACATTTA	1956
Db		1721	CGCTGTGGGAACACCGGAGAGCTGTGTCTTTTGAACCCGGGGGCTGACGGCAACTGT	1780
Qy		1957	GCACAGCCCTCTAGTAAAGTATTGACAGANTCTCAACAGATGTTCCTGCATGTTGGAA	2016
Db		1781	GCGGCTCCCTCGTCAAGTCTGTGTCTGGGGCCGACGAGAGAGTGTCCCAATATGCTGGAG	1840
Qy		2017	GAATTTGCTTTTACTGAC	2033
Db		1841	GAACTGCTTTCAGTGC	1857
<hr/>				
RESULT_14				
AF479823				
LOCUS		1512 bp	mRNA	linear VRT 01-APR-2002
DEFINITION	Pantodon buchholzi	DEAD box RNA	helicase Vasa mRNA,	complete cds.
ACCESSION	AF479823			
VERSION	AF479823.1	GI:19387222		
KEYWORDS				
SOURCE				
ORGANISM	Pantodon buchholzi.			
	Pantodon buchholzi			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Actinopterygii; Neopterygii; Telosteii; Osteoglossomorpha;			
	Osteoglossiformes; Pantodontidae; Pantodon.			
REFERENCE	1 (bases 1 to 1512)			
AUTHORS	Knaut,H., Steinelsner,H., Schwarz,H. and Nusslein-Volhard,C.			
TITLE	An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA			
JOURNAL	Translaction to the Germ Cells in the Zebrafish			
MEDLINE	Curr. Biol. 12 (6), 454-466 (2002)			
PUBMED	21906633			
REFERENCE	11909530			
AUTHORS	2 (bases 1 to 1512)			
TITLE	Knaut,H.			
JOURNAL	Direct Submission			
FEATURES	Submitted (04-FEB-2002) Genetik, Max-Planck Institut fuer			
	Entwicklungsbiologie, Spemannstrasse 35, Tuebingen 72076, Germany			
	Location/Qualifiers			
	1..1512			

		/organism="Pantodon buchholzi"
	/db_xref="taxon:8276"	
CDS	1..1185	
	/codon_start=1	
	/product="DEAD box RNA helicase Vasa"	
	/protein_id="AL87142.1"	
	/db_xref="GI:19387223"	
	/translation="MACAQTGSGCTAAFLPILOQMTGSVASSFESEVEPEAIIVA PYEELINQIYLEAEKFAFGTCVRPVVYGSTMYTIREVLKGNILCGTGRLDII GRGVNGIKYFLEVLDEADMIDMPEMDRRLIVASGMPEKERQLMFATPEDP ORLADPLAKRDYFLVVGVGAGCNDVGEHVSYSKEOLVEIJRTTSNETVMV VERTRKADFIATFLFCOISITSSIGHGREOREBOALGDPRSGCPVLAVTSVARG DIHVHQVNVNFDLPSTIDEXYHRIGRGRCGNTKAISFPDPEDAPLASLYKLSS AQGVPMLEEIASAHGTAFNPGRGVASTDRKEGAPRAEPVPAVNPDADER WE"	
BASE COUNT	387 a 301 c 406 g 415 t	3 others
ORIGIN		
Query Match	26.0%; Score 565; DB 5; Length 1512;	
Best Local Similarity	71.0%; Pred. No. 8.9e-123;	
Matches 745; Conservative	2; Mismatches 302; Indels 0; Gaps 0;	
Oy	985 ATGGCTGTGCTCAACAGGGTCTGGGAAGACTGCGCTTTTCCCTAACAAATTGGACT	1044
Dd	1 ATGGCATGCTCTACACTGCATCGGGTAAGCGCTGATCTTTATTAACCATCTTGCAA	60
Oy	1045 CATATGATGATGATGAAATACTGCCAGTCGTTTTAAAGATTGCAGGAACAAGATGT	1104
Dd	61 CAGCTGATGACACAGCGGTGTGGCTTCCAGTCGGTTTAGTGAAGTCAGAGACCGGAGCC	120
Oy	1105 ATTATTTGTAGCACCACTCGAGAATTGGTCAACAGATTATTTTGGAAACCCAGAAAATT	1164
Dd	121 ATTGTGTGGCCCCAACAGAGACTTATCAACAGATATATCTGGAAGCCAGAGAAATT	180
Oy	1165 TCTTTGGGACTGTGTGAAGAGCTGTTGTTATATATGGGGGAACCCAGCTGGGACATTCA	1224
Dd	181 GCCTTTGGAGCTGTGTGAGCGCTGTGTGTGTATGAGAGAACCTGACCATGTACAC	240
Oy	1225 ATTGCACAAATAGTACACAGCGCTGTAAATATATATAGTCTACTCCGTGAAGACTATGAT	1284
Dd	241 ATACGTGAAGTTTGAAGAAGCTGCAACATTTATAGTGGACACCTGGAGACTGCTTGAC	300
Oy	1285 ATCATAGGCAAGAAAAAGATTGCTCAAAACAGATCAAAATACCTTAGTTTGGATGAAGCT	1344
Dd	301 ATTATGGAAGAGAAAGTTGGCGCTGAATTAAGCTTCGTACCTTGTAAGTGAAGGCA	360
Oy	1345 GATGCGATGTTGGATATAGGCTTTTGGTCCAGAAATGAAGAAGTTAATTTCTTCCCACAGA	1404
Dd	361 GACCGCATGTTGGATATAGGCTTTTGAACCAAGACATGGCGAGGTGGCTTCTCCAGGC	420
Oy	1405 ATGCGATCAAAAGAACACGCGCCAACCCCTTATATGTTCAAGTTCAGAGGAATTT	1464
Dd	421 ATGCGCTCCMAAGAGGAGCCACAGACACTTATATGTTCACTGACCTACCCAGAGACATT	480
Oy	1465 CAAGGTTGGCTGCAGAGTTTAAAGTCAAATATCTGTTGTGCTGTGACAAAGTG	1524
Dd	481 CAAAGGCTGCTGCAGAGTTTCTTGAAGAGACTACTGTTCTTAAGTTGTGGCAAGTG	540
Oy	1525 GGTGGAGCATGTAGAGATGTTACAGACACCGCTTCCAAAGTTGGCCAGTTTCTCAAAAAGA	1584
Dd	541 GGTGGAGCTGCAGATGATGTTGAACACACAGTATGTAAGTGTCCAGTATTCAAAGCG	600
Oy	1585 GAAGAAGCTCGTGAATTTCTGCGAACAATAGGGGATGAAGAATAAGTCTTTTGTGGA	1644
Dd	601 GAGACGCTGTGTGAGATCTTCAGAACTACAGGAATAAGCGCACTATGTGTGTGTGGA	660
Oy	1645 ACTAAGAAAAAGCAGATTTTACTGCACTTTTCTTTGTCAAGAAAAAATATCAACTACA	1704
Dd	661 ACCAAAGAAAGGAGAGATTTTATGCGCACTTTCTCTGTGAGAAACAGATATCAACACA	720
Oy	1705 AGTATTCATGTGATTCGGGAACAGAGAGCGGGAGCAAGCTCTTGGAGATTTTTCGCTTT	1764
Dd	721 AGTATTCATGTGATTCGGGAACAGAGAGCGGTAGCAAGGCTTGAAGTGTGATCTTCGCGAGT	780

QY 1765 GGAAGTCCAGTCTTCTTGTGCTACTTCACTAGTCTGCCAGAGGGCTGATTTGAAAT 1824  
 DB 781 GGGAAATGCTCTCTCTTGTGGCCACTCGGTGGACGACAGGTCTTGACATTGACAT 840  
 QY 1825 GTGCAACATGTTATCAATTTTGAATCTTCTTACCATGATGATAATGTTACGAAT 1884  
 DB 841 GTCCAGCATGTGGTGAATTTGACCTCCGTCACCATTTGATGATAATGACCCGAT 900  
 QY 1885 GGGCGTACTGCTGCTTGTGGAACTCTGCAAGCAATTTCTTTTGTGATCTTGAATCG 1944  
 DB 901 GGGGGGACAGGCGCTGTGGAACTCTGCAAGCAATTTCTTTTGTGATCTTGAATCG 960  
 QY 1945 GATACACATTTGACAGCCCTCTAGTAAAGTATGACAGATGCTCAAGAGATGTTCT 2004  
 DB 961 GATGCCCACTTGTCTCTCTCTGTTAAAGTTCTTTCAAGGGCCGACAGAAAGTTCC 1020  
 QY 2005 GCATGTTGAGAAATTTGCTTTAGTAC 2033  
 DB 1021 AAGTGGCTGAAGAAATTTGATTTAGTGC 1049  
 RESULT 15  
 AB063484 2187 bp mRNA linear VRT 22-JUN-2001  
 LOCUS Oryzias latipes olvas mRNA for VASA, complete cds.  
 DEFINITION AB063484  
 ACCESSION AB063484 GI:14522856  
 VERSION AB063484.1 GI:14522856  
 KEYWORDS Oryzias latipes cDNA to mRNA.  
 SOURCE Oryzias latipes  
 ORGANISM Oryzias latipes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.  
 1 (sites)  
 Shihomiya, A., Tanaka, M., Kobayashi, T., Nagahama, Y. and Hamaguchi, S.  
 The vasal-like gene, olvas, identifies the migration path of primordial germ cells during embryonic body formation stage in the medaka, *Oryzias latipes*  
 Development 42, 317-326 (2000)  
 2 (sites)  
 Tanaka, M., Kinoshita, M., Kobayashi, D. and Nagahama, Y.  
 Establishment of medaka (*Oryzias latipes*) transgenic lines with the expression of green fluorescent protein fluorescence exclusively in germ cells: a useful model to monitor germ cells in a live vertebrate  
 Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2544-2549 (2001)  
 JOURNAL MEDLINE 21126920  
 PUBMED 11226275  
 REFERENCE 3 (bases 1 to 2187)  
 AUTHORS Tanaka, M.  
 JOURNAL Direct Submission  
 TITLE Submitted (20-JUN-2001) Minoru Tanaka, Hokkaido university, Graduate school of Science, division of Biological science, N10, W8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan  
 (E-mail: mtanaka@sci.hokudai.ac.jp, Tel: 81-11-706-4460 (ex. 4460), Fax: 81-11-706-4460)  
 FEATURES  
 source location/Qualifiers  
 1. 2187  
 /organism="Oryzias latipes"  
 /db\_xref="taxon:8090"  
 /note="Japanese killifish, medaka-himedaka"  
 1. 2187  
 /gene="Olvas"  
 94. 1947  
 /gene="Olvas"  
 /codon\_start=1  
 /product="VASA"  
 /protein\_id="BAB1047.1"  
 /db\_xref="GI:14522857"  
 /translation="MDWEETAPSFAPVSTDAFORSMNGARDSDGNDGSMNR SNRGGSAGRGGRGRGRGRSDDELNGGSDSNGFRGRGRGRGRSGR  
 CDS  
 gene  
 CDS

GGGVGRGDEDFVAGDORGANSDAADPERKVTYIPSPLEDEDSIFSHYKGIN  
 PKYDDIIVDVSGTALPAIMTFEAKACESENNIISGIVKPVQVYGLPIISAG  
 RDLMAKDEKTAARFLPILOQMAKDVRSREIOPAVPVAIPRELINOIYO  
 EARKSRPCVCPVYVGVGTCGPGMGLMIGKXGLSVR  
 HLVLBADMDMGFEEDPDKRLVSGFPCMRSEKREOTLWFSATFPEDIQBLADFLKVD  
 YLFVAVGVGACTVEQTFLOVTFENKREQLDLRTIGSERTEWFEYKQADDTIA  
 AFLCOEKVETTSIHDRBQREKELADPFRSKCVLVAIVASRGLDIPVOHVNF  
 DLPTIDIVYHRIGRGRCNTGRAVSPFDPPVDLSQARSLVIGILAKAOQEVPSMLEE  
 SAFGAGSAAFNPSSRTFASVTSRKGSGQDSVXTOPAPAAPADEDDME"  
 BASE COUNT 489 a 610 c 715 g 373 t  
 ORIGIN  
 Query Match 25.5%; Score 552.8; DB 5; Length 2187;  
 Best Local Similarity 61.8%; Pred. No. 6.7e-120;  
 Matches 897; Conservative 0; Mismatches 552; Indels 3; Gaps 1;  
 QY 711 TGATTAACAAGACCAAAAGTGAACCTTACCTCCCTCCACCTGAGAGATGAGACTC 770  
 DB 483 TGACCCAGAACGCGCCAAAGTGAACCTTACCTCCCTCCACCTGAGAGAGACTC 542  
 QY 771 CATCTTGACATTTATCAGACAGGACATTAACCTTCGACAAATAGACACTATTTCTTGG 830  
 DB 543 CATCTTCCCACTACAAAGATGGGCACTTTCAGCAAGTACAGACATCTGTGTGA 602  
 QY 831 AGTGTCTGACATGATGACCAACACCAATTCGATTTTGAAGAAGCTAATCTGTGA 890  
 DB 603 CGTCAAGGGGACCAACTCGCGCGCATTCATGACTTTGAGAGAGCCAGCTGTGCGA 662  
 QY 891 GACACTGAATTAACACATTTGCTAAGCTGTATTAAGCTTTACTCTCTGTGCAAAATA 950  
 DB 663 GTGCTTGAGAAACAAATCAGACAGGTCCGAGTACGTCAAGCCGACCCGTCAGAAAGA 722  
 QY 951 CAGTATTCATCATCTTGAAGAGAGATTTGATGGCTGTGCTCAACAGAGCTGTGA 1010  
 DB 723 CGGCTCCCATATCTCCGCGGAGAACTGATGGCTGTGCTCAACAGAGCTGTGA 782  
 QY 1011 GAAGACTGGGGCTTTTCTCTCAATTTTGGCTCATATGATGATGATGAATTAATGCTGC 1070  
 DB 783 AAAAATGGGGGCTTCTCTGCTGCATCTGACAGAGCTGATGGAGACGGCTGGCGGC 842  
 QY 1071 CAGTCTTTTAAAGATTGACAGAACCAAGCTATTAATTTAGCAACCACTGAGAAAT 1130  
 DB 843 CAGCCCTTCAGGAGATCCAGAGACAGAGCCGATTTGGCTTCAACAGGAGAGCT 902  
 QY 1131 GGTCAACGAGTTTATTTGGAAGCCAGAAATTTCTTTTGGAGCTTGTAGAGACTGT 1190  
 DB 903 CATCAACGAGTTTATTCAGAGAGCCCGGAGATTTCTTTGGACCTGCTGGCGCCGT 962  
 QY 1191 TGTATATATGGGGGAACCCAGCTGGACATTCATTCGACAAATAGTCAAGGCTGTAA 1250  
 DB 963 GGTGTCTACGAGGGGTGAACACCGGCTACAGATGAGGAGATCGAGAAAGGCTGCAA 1022  
 QY 1251 TATATTATGTCTACTCTCGGAAGACTGATGATATCATATAGCCAAAGAAAGATTGCT 1310  
 DB 1023 CGTGTCTGGGAGACTCCCGGAGCCTGTGTGCAATGATCGGAGAGAAAGGTGGCCT 1082  
 QY 1311 CAACAGATTAATTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1370  
 DB 1083 GAGTAAAGTTCGACCTGTCTCTGACGAGCGACCGCATGTGACATGGGCTTCCA 1142  
 QY 1371 TCCAGAAATGAAGATTAATTTCTTGCCAGAGATGCCATCAAGAGAACGCGCAAC 1430  
 DB 1143 GCGGACATCGCGCGCTGTGGGCTCCCGGCTATAGCGCTCCAAAGAGAGGGCAGAC 1202  
 QY 1431 CTTATGTTCAAGCACTTTTCCAGAGAAATTCAAAGGTTGGCTGACAGATTTTAA 1490  
 DB 1203 GGTGATGTTCAAGCACTTTCCCGAGGACATCCAAAGGTTGGCGCGCACTTCTGAA 1262  
 QY 1491 GTCAATTAATCTTTGTTGCTGTTGCAAGAGGGGTGAGCATGTAGATGTTGCA 1550  
 DB 1263 GTTGAATCACTTTTGTGGCTGTGGGGGTGTGGGGGGAGCTTGCACCGAGCTGAGCA 1322  
 QY 1551 GACCGTTCTCAAGTTGGCCAGTTCTCAAAAAGAGAAAGCTGTTGAAATTTCTGCGAAA 1610

```
Db      1323  GACCTTCCTACAGTCAACCAATTCAACAAGCGGAGCAGCTCTGACCTCCTCAGGAC 1382
QY      1611  CATAGGGGATGAAGAAGCTATGGCTTTGTTGAACCTAAGAAAAAGCAGATTTTACTGC 1670
Db      1383  CATGGGCTCGAGCGGACATGGTGTTCGTGAGACCCAGAGGCGAGCGACTTTCATCGC 1442
QY      1671  AACTTTCTTTGTCAGAAAAAATATCACTACAGATATCATGATCGGATCGGAGCAGAG 1730
Db      1443  CGCCTTCCTGTGCAAGAGAGGTTCGACACACAGCATCCAGGCGACAGAGAGCAGCG 1502
QY      1731  AGAGCGGAGCAAGCTCTTGGAGATTTGCTTGGAAAGTCCCAAGTTCTTGTGCTAC 1790
Db      1503  CGAGCGGAGAGAGCGCTGGCGGATTTCCGCTCAGGCAAGTGTCCGCTCTGGTGGCCAC 1562
QY      1791  TTCAGTAGCTGCAGAGGGGCTGGATATGAAATGTCAACATGTTATCAATTTGATCT 1850
Db      1563  CTGGTGGCGTCCGCGGCTGGACATCCTGACGTCCAGACAGTGTGTAATTTGACTT 1622
QY      1851  TCCTTACCATGATGATATATGTTTCATCGAATGGGCGTACTGTCGTTGGGAATAC 1910
Db      1623  GCCCAACACCATTTGAGACTAGCTCCACCGCATCGGAGAGACGGGCGCTGCCGCAACAC 1682
QY      1911  TGGCAGAGCAATTTCTTTTGTGATCTTGATCGGATACCATTTAGCACAGCCTTAGT 1970
Db      1683  CGGCAAGGCGGTGTCTTCTTACGACCCCGACGTGACAGTCACTGGCCCGCTCGCTGCT 1742
QY      1971  AAAAGTATTGACAGATGCTCAACAGAGATGTCTGCAATGTTGGAGAAATTCCTTAG 2030
Db      1743  CGGCATCTCTGGCAAGGCTGACAGAGAGTGCCTCATGGCTGAGAGATCGGCGTTGCG 1802
QY      2031  TACATACATTCTGCTTCAAGTAGTACAGAGAGAAAGTGTTCATCAGTTGATAC 2090
Db      1803  CGCTCAAGGTTCCGCCGCC--TTCAACCTTCGGGAGAGACGTTTGCTTCCACAGACTC 1859
QY      2091  CAGAAAGGCAAGAGCACTTGAACACAGCTGGGTTTCTTTCACAGAGCTCCCAATCC 2150
Db      1860  CAGGAGAGGTGTCTTTCAGAGCAGCAGTGTGAAGACGACGCTGACGCCCCCTGC 1919
QY      2151  AGTAGATGATGA 2162
Db      1920  TGCTGTGATGA 1931
```

Search completed: June 10, 2003, 09:43:34  
Job time : 5469 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 14:53:30 ; Search time 28 Seconds  
(without alignments)  
6434.756 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 3954  
Sequence: 1 atgggggagatgaagatcggga.....tagatgatgagtcagtcgggat 2172

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 6.0, Ygapext 7.0  
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p model -DEV=xld  
-Q=/cgn2\_1/USPRO.spool/US09714865/runat\_05062003 11758\_26006/arp\_query\_faeta.1.2311  
-DB=SwissProt 40 -QFMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09714865 @CGN 1.1 23 @runat\_05062003 11758\_26006 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WANT\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3813	96.4	724	DDX4_HUMAN	O9NQ10 homo sapien
2	3249	82.2	702	DDX4_MOUSE	O61496 mus musculu
3	3247.5	82.1	713	DDX4_RAT	O64060 rattus norv
4	1290	32.6	661	VASA_DROME	P09052 drosophila
5	1229	31.0	661	DDX3_MOUSE	O62167 mus musculu
6	1224.5	31.0	697	AN3_XENLA	P24346 xenopus lae
7	1215	30.7	661	DDX3_HUMAN	O00571 homo sapien
8	1197	30.3	660	P110_MOUSE	P16381 mus musculu
9	1182.5	29.9	660	DDX1_HUMAN	O15523 homo sapien
10	1126.5	28.5	604	DDX1_YEAST	P06634 saccharomyc
11	1109	28.0	636	DDX1_SCHPO	O13370 schizosacch
12	1108	28.0	617	DDX1_YEAST	P24784 saccharomyc
13	1096.5	27.7	763	GLH1_CAEEL	P34689 caenorhabdi
14	992.5	25.1	720	GLH3_CAEEL	O01836 caenorhabdi
15	820.5	20.8	550	DBP2_SCHRO	P24782 schizosacch
16	815	20.6	546	DBP2_YEAST	P24783 saccharomyc
17	795.5	20.1	575	RM62_DROME	P19109 drosophila
18	777.5	19.7	614	DDX5_HUMAN	P17844 homo sapien

19	777.5	19.7	614	DDX5_MOUSE	O61656 mus musculu
20	771.5	19.5	1156	GLH4_CAEEL	O67443 caenorhabdi
21	754	19.1	607	DB10_NICSY	P46942 nicotiana s
22	754	19.1	650	DD17_HUMAN	O92841 homo sapien
23	748	18.9	568	HE64_TRYEB	O26696 trypanosoma
24	693.5	17.5	523	DBP3_YEAST	P20447 saccharomyc
25	692.5	17.5	578	YBX6_SCHPO	Q10202 schizosacch
26	685	17.3	454	RHLE_ECOLI	P25888 escherichia
27	684	17.3	245	DEA2_MOUSE	O62095 mus musculu
28	675	17.1	622	DEA2_HUMAN	O91979 mus sapien
29	668.5	16.9	619	ABS_DROME	O9130 drosophila
30	621	15.7	524	EXP5_STRPN	P35599 streptococc
31	594	15.0	533	RRP3_YEAST	P38712 saccharomyc
32	587	14.8	588	PR28_YEAST	P23394 saccharomyc
33	575.5	14.6	481	YAB8_SCHPO	O09775 schizosacch
34	569	14.4	563	DEAD_MCTU	O11039 mycobacteri
35	567.5	14.4	642	DEAD_KLEPN	P33906 klebsiella
36	566.5	14.3	628	DEAD_ECOLI	P23304 escherichia
37	563.5	14.3	851	DD21_MOUSE	O91145 mus musculu
38	562.5	14.2	613	DEAD_HAEIN	P44586 haemophilus
39	560	14.2	715	DD21_HUMAN	O9130 homo sapien
40	556	14.1	601	DEAD_BUCAI	P57453 buchnera ap
41	554.5	14.0	848	YA47_SCHPO	O09719 schizosacch
42	548.5	13.9	564	R0K1_YEAST	P45818 saccharomyc
43	546.5	13.8	479	DBPA_BACSU	P42305 bacillus su
44	542	13.7	457	DBPA_ECOLI	P21693 escherichia
45	533	13.5	754	YAJ3_SCHPO	O09903 schizosacch

## ALIGNMENTS

RESULT 1  
ID DDX4\_HUMAN STANDARD; PRT; 724 AA.  
AC O9NQ10; O9NQ10; O9NT92;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE DEAD-box protein 4 (VASA homolog).  
GN DDX4 OR VASA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
RC TISSUE=Testis;  
RA MEDLINE=20402578; PubMed=10920202;  
RX Castrillon D.H., Quade B.J., Wang T.Y., Outgley C., Crum C.P.;  
RT "The human VASA gene is specifically expressed in the germ cell  
lineage";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:9585-9590(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Rocha D., Affara N.;  
RT "Cloning and characterization of the human VASA gene";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 90-724 FROM N.A.  
RC TISSUE=Testis;  
RA Blum H., Bauersachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;  
RT TISSUE=Testis;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: May play a role in germ cell development.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- TISSUE SPECIFICITY: Expressed only in ovary and testis. Expressed  
CC in migratory primordial germ cells in the region of the gonadal  
CC ridge in both sexes.  
CC -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
CC DDX4/VASA SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: AY004154; AAF6585.1; -  
 DR EMBL: AF262962; AAF72705.1; -  
 DR EMBL: AL137462; CAB70750.1; -  
 DR HSSP: Q56083; IHV8.  
 DR GeneW: HGNC:18700; DDX4.  
 DR MIM: 605281; -  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR000629; DEAD box.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR Pfam: PF00270; DEAD. 2.  
 DR Pfam: PF00271; Helicase\_C; 2.  
 DR SMART: SM00487; DEXDC; 1.  
 DR SMART: SM00490; HELIC; 1.  
 DR PROSITE: PS00039; DEAD ATP HELICASE; 1.  
 DR Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.  
 KW NP BIND 332 339 ATP (PORENTIAL).  
 FT SITE 446 449 DEAD BOX.  
 FT DOMAIN 58 234 GLY-RICH.  
 FT CONFLICT 556 556 I -> T (IN REF. 1).  
 FT CONFLICT 713 713 Q -> R (IN REF. 1).  
 SQ SEQUENCE 724 AA; 79307 MW; 6D313DD98B177067 CRC64;

## Alignment Scores:

Score: 1.57e-261 Length: 724  
 Percent Similarity: 3813.00 Matches: 722  
 Best Local Similarity: 99.86% Conservative: 1  
 Query Match: 96.43% Mismatches: 1  
 DB: 1 Indels: 0  
 Gaps: 0

US-09-714-865-15 (1-2172) x DDX4\_HUMAN (1-724)

QY 1 ATGGGGGATGAAGATTGGGAGAGCAAGATCAACCTCATATGCTCTCTATGTTCCATA 60  
 Db 1 MetGlyAspGluAspTrpGluAspIleAspProHisMetSerTyrValProIle 20  
 QY 61 TTGAGAGGATAGGATATTCTGAGAGAAATGAGACATTTTAAACAGGACTCCAGCTTCA 120  
 Db 21 PheGluAspAspArgTyrSerGlyGluAsnGlyAspAsnPhenAlaArgThrProAlaSer 40  
 QY 121 TCATCAGAAATGATGATGACCTTCTGAAAGATCTATTCAATGAAAAGTGATTTGCC 180  
 Db 41 SerSerGluMetAspAspGlyProSerArgArgAspHisPheMetLysSerGlyPheAla 60  
 QY 181 TCTGGGCGGAAATTTTGAAGACAGAGATGCTGAGTAAATAAGCCGATATACATCC 240  
 Db 61 SerGlyAspAsnPheGlyAsnArgAspAlaGlyGluCysAsnLysArgAspAsnThrSer 80  
 QY 241 ACAATGGGTGTTTGGAGTTGAGAAAGATTGGAAACAGAGTTTTCAAACAGCAGG 300  
 Db 81 ThrMetGlyGlyPheGlyValGlyLysSerPheGlyAsnArgGlyPheSerAsnSerArg 100  
 QY 301 TTTGAAATGCTGATAGCTCTGCTTCTGAGAGAGCTTAAATGATGCTGGAGATTAAT 360  
 Db 101 PheGluAspGlyAspSerSerGlyPheThrArgLysSerSerAsnAspCysGluAspAsn 120  
 QY 361 CCAACACGGAACAGAGGGTTTTCAGAGAGCGGCTATCGAGATGAGAAATTAATTCAGAA 420  
 Db 121 ProThrArgAsnArgGlyPheSerLysArgGlyLysTrpArgAspGlyAsnAsnSerGlu 140  
 QY 421 GCTTCAGGCGCATACAGAAAGGTGGAAGAGTAGTATTCAGAGTTGCCGTGAGAGATT 480  
 Db 141 AlaSerGlyProTyrArgArgGlyGlyLysArgSerPheArgGlyCysArgGlyGlyPhe 160  
 QY 481 GGTCTGAGAGTCCAAATTAATGACTTAGACCCAGACGAAATGTATGACAGCGACTGGTGGC 540  
 Db 161 GlyLeuGlySerProAsnAsnAspLeuAspProAspGluCysMetGlnArgThrGlyGly 180

QY 541 CTTTGTGTTCTAGAGACAGCATATTAGTGACAGAGTAATGATGATCTTCAAGC 600  
 Db 181 LeuPheGlySerThrArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSer 200  
 QY 601 AGAAGTGGCAGTGGAGTGAACAGAGGTGTTACAAAGCTTTAAATGAAGATTAACA 660  
 Db 201 ArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThr 220  
 QY 661 GGGCTGGAAAGAAATTTCTGGAAGTCCAGAGCAAGAGAGAGAGAAAGTAGTACTCAA 720  
 Db 221 GlySerGlyLysAsnSerTrpLysSerGluAlaGluGlyGlyLysSerSerPheGln 240  
 QY 721 GGAACAAAGTGAACCTCATACCCCTCTCTCACTGAGAGTAGAGACTCCATCTTTGCA 780  
 Db 241 GlyProLysValThrTyrIleProProProProProGluAspGluAspSerIlePheAla 260  
 QY 781 CATTATCAGACAGGCATTAATCTTGACAAATACGACACTATTCTTGGAAGTCTGGA 840  
 Db 261 HisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSerGly 280  
 QY 841 CATGATGCACACACGCAATTTGACTTTGAGAGACTAATCTCTGTCAGACCTGAT 900  
 Db 281 HisAspLysProProAlaIleLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsn 300  
 QY 901 AACACATTTGCTAAAGCTGGTATATCTAATGACTTACTCTCTGCAAAAATACAGATTCT 960  
 Db 301 AsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIlePro 320  
 QY 961 ATCATCTTCGACAGACAGATTTGATGCTTGCTCAAAACAGGCTTGGGAAGACTGGC 1020  
 Db 321 IleIleLeuAlaGlyArgAspLeuMetAlaCysIleGlnThrGlySerGlyLysThrAla 340  
 QY 1021 GCTTTTCTTACCAATTTTGGCTCATATGATGATGATGATGATGATGATGATGATGAT 1080  
 Db 341 AlaPheLeuLeuProIleLeuAlaHisMetMetHisAspGlyIleThrAlaSerArgPhe 360  
 QY 1081 AAAGAGTTGACAGAAACAGAGTATATTATGTAGCACCAATCGAGATTGGTCAACAG 1140  
 Db 361 LysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuValAsnGln 380  
 QY 1141 ATTATTTGGAAGCCAGAAAATTTCTTTTGGGACTTGTGTAAGCTGTTGTTATATAT 1200  
 Db 381 IleTyrLeuGlnAlaArgLysPheSerPheGlyThrCysValArgAlaValIleTyr 400  
 QY 1201 GGGGGAACCCAGCTGGGACATTCATTCGACCAATATGACAGGCTGTAATATATAT 1260  
 Db 401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlyCysAsnIleLeuCys 420  
 QY 1261 GCTACTCTCGAAGACTGATGATATCATAGGCAAGAAAGATTTGCTCAACAGATC 1320  
 Db 421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGlyLysIleGlyLeuLysGlnIle 440  
 QY 1321 AAATACTTGAATTTGATGAAGCTGATGATGATGATGATGATGATGATGATGATGAT 1380  
 Db 441 LysTyrLeuValIleAspGluAlaAspArgMetLeuAspMetGlyPheGlyProGluMet 460  
 QY 1381 AAGAAGTTAATTTCTGCGCAGGAATGCCATCAAGGAAGAGCGCCAAACCTTAATGTC 1440  
 Db 461 LysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeuMetPhe 480  
 QY 1441 AGTGCACTTTTCAGAGAAATTCAGAGTTGCTGACAGATTTTAAAGTCAAAATTAAT 1500  
 Db 481 SerIleAlaThrPheProGluGluIleGlnArgLeuAlaIleGluPheLeuLysSerAsnTyr 500  
 QY 1501 CTGTTTGTGCTGTGGAACAAGTGGTGAAGCATGTAGAATGTTCCAGCAGACCTGTTCTC 1560  
 Db 501 LeuPheValAlaValGlyGlnValGlyValAlaCysArgAspValGlnGlnThrValLeu 520  
 QY 1561 CAAGTGGCAGTCTCAAAAAGAGAAAGCTGTTGAATTTTCGGAACATAGGGGAT 1620  
 Db 521 GlnValGlyGlnPheSerLysArgGluLysLeuValGluIleLeuArgAsnIleGlyAsp 540  
 QY 1621 GAAAGAACTATGCTTTGTTGAATAAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680

Db	Accession	Gene	Protein	Length
Dp	541	Glutathione S-transferase P1	560	560
Qy	1681	TGTCAGAAAAAATATCAACTACAGATATTCATGGTGTATCGGAAACAGAGAGACGGAGG	1740	1740
Dp	561	CyegingiluvslieserThrThrsrllshsglsaparglungrglungrglu	580	580
Qy	1741	CAAGCTCTTGGAGATTTTCGCTTTGGAAAGGCCAGCTTTCTTGTCTTACTTCACTTCACT	1800	1800
Dp	581	GlnalaleugglyaspPheargPheclLysCysProValIeuaIalathSerValala	600	600
Qy	1801	GCCAGAGGGCGGATATTTGAATATGTCAGACATGTTATCAATTTTGATCTTCCTTACAC	1866	1866
Dp	601	AlaargglyleuasplleglunsnvalglnhsvallleasnpeasplneproserThr	620	620
Qy	1861	ATTGATGAATATGTTTCATCGAATTTGGGCGTACTGGCTTGTGGAAATATTCAGAGACA	1920	1920
Dp	621	IleaspluvryvalHlsarlglllelyarghrgrlgaycysglYasnThrGlyArgsla	640	640
Qy	1921	ATTTCTCTTTTGGATCTTGAATCGGATACCATTTAGACACGCTCTTGTCTTAAAGTATG	1980	1980
Dp	641	IleaserPhepPheasplleuglLyserspsnhsleualaglnProleuVallyValleu	660	660
Qy	1981	AACAGTGTCTCAACAGGAGTTCCTGTCAGATGTTGGAAAGAAATTCCTTAGTACATACAT	2040	2040
Dp	661	ThrspslaaglnglnasplvalProalatrleuglungrluallealapheserThrlylle	680	680
Qy	2041	CCTGGCTTCAGTGTGATGATCAAGAGAAACGTTTGCATGCATGATGATACCAAGAAAGGC	2100	2100
Dp	681	ProglPheaserGlyserThrArglyasnvalPhealaserValasplThrArglyselly	700	700
Qy	2101	AAGACGACTTTGAACACAGCTGGGTTTCTTCTTTCACGAGCTCCCAATCCAGTATGAT	2166	2166
Dp	701	LysserThrleuasThrAlaglyPheaserSerSerGlnalalProasProValasplasp	720	720
Qy	2161	GAGTCATGGGAT 2172		
Dp	721	GlusertPasp 724		

```

RA  Havaishizaki Y.:
RT  "Functional annotation of a full-length mouse cDNA collection."
RL  Nature 409:685-690(2001).
RN  [2]
RP  SEQUENCE OF 60-702 FROM N.A.
RX  STRAIN=BALB/c; TISSUE=Testis;
RA  MEDLINE=95083681; PubMed=7991615.
RA  Fujiwara Y., Komiya T., Kawabata H., Sato M., Fujimoto H.,
RT  Furusawa M., Noce T.;
RT  "Isolation of a DEAD-family protein gene that encodes a murine
RT  homolog of Drosophila vasa and its specific expression in germ cell
RT  lineage.";
RL  Proc. Natl. Acad. Sci. U.S.A. 91:12258-12262(1994).
CC  -1- FUNCTION: May play a role in germ cell development.
CC  -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR.
CC  -1- TISSUE SPECIFICITY: Testis.
CC  -1- DEVELOPMENTAL STAGE: EXPRESSED IN SPERMATOGENIC CELLS FROM THE
CC  SPERMATOCTYE STAGE TO THE ROUND SPERMATID STAGE.
CC  -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC  DDX4/VASA SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AK014844; BAB29578.1; -.
DR  EMBL; D14859; BAA03584.1; -.
DR  HSSP; Q58083; IHV8.
DR  MGD; MGI:102670; Ddx4.
DR  InterPro; IPR001410; DEAD.
DR  InterPro; IPR000629; DEAD box.
DR  InterPro; IPR001650; Helicase_C.
DR  Pfam; PF00270; DEAD_2.
DR  Pfam; PF00271; helicase_C_2.
DR  SMART; SM00487; DEXDC; 1.
DR  SMART; SM00490; HELICC; 1.
DR  PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW  Developmental protein, Hydrolase, ATP-binding, Repeat, Helicase.
KM  NP_BIND 305 312
FT  SITE 419 422
FT  DOMAIN 58 207
FT  CONFLICT 152 152 R->C (IN REF. 2).
FT  CONFLICT 156 160 LFGR -> FLVIG (IN REF. 2).
FT  CONFLICT 277 277 A->R (IN REF. 2).
FT  CONFLICT 291 291 S->T (IN REF. 2).
FT  CONFLICT 368 369 RA->IS (IN REF. 2).
FT  CONFLICT 423 423 R->S (IN REF. 2).
FT  CONFLICT 430 430 G->A (IN REF. 2).
FT  CONFLICT 433 433 M->I (IN REF. 2).
FT  CONFLICT 448 448 R->H (IN REF. 2).
FT  CONFLICT 472 472 S->N (IN REF. 2).
FT  CONFLICT 499 502 SKRKLVEILR -> QREKSLREYE (IN REF. 2).
FT  CONFLICT 540 540 T -> S (IN REF. 2).
FT  CONFLICT 540 540 HTLTAGISSQAQNPVDDSDMD -> AHVEYSGDFPFTSS
FT  CONFLICT 702 702 QSS (IN REF. 2)
SQ  SEQUENCE 702 AA; 76470 MW; 5D1E4CC0BDF9748 CRC64;

Alignment Scores:
Pred. No.: 1.01e-221 Length: 702
Score: 3249.00 Matches: 626
Percent Similarity: 90.42% Conservative: 35
Best Local Similarity: 85.64% Mismatches: 34
Query Match: 82.17% Indels: 36
DB: 1 Gaps: 7

US-09-714-865-15 (1-2172) x DDX4_MOUSE (1-702)
1 ATGGGGATGAAGATTGGGAAGCAATC---AACCTCATATGCTTCTTATGTTCC 57

```

Db 1 MetGlyAspGluAspTrpGluAlaGluIleuLeuysProHisValSerSerTyValPro 20  
 QY 58 ATATTGAGAAAGATAGTAT---TCTGGAGAAAAATGAGACATTTTAAACAGACCTCCA 114  
 Db 21 ValPheGluIysAspLysTySerSerGlyAlaAsnGlyAspThrPheAsnArgThSer 40  
 QY 115 GCTTCATCATCAAGAAATGATGATGACCTTCGAAAGATCATTTTCATGAAAAAGTGA 174  
 Db 41 AlaSerSer----- 43  
 QY 175 TTTGCTCTGGCGGAAATTTTGAACAGAGATCTGTGATGTAAATAGCGAGATPAT 234  
 Db 44 -----AspIleGlyIuSerSerTyLysGluAsn 53  
 QY 235 ACATCCACAAATGGGTGTTTGGAGTTTGGAAAAGATTGGAAACAGAGCTTTTCAAC 294  
 Db 54 ThrSerThrThrGlyGlyPheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAsn 73  
 QY 295 AGCAGGTTTGAAGATGGATAGCTGTGTTTCTGAGAGAGAGTCTAGTAATGATGCGAA 354  
 Db 74 AsnLysPheGluGluGlyLysPheSerSerGlyPheTrpLysGluSerAsnAsnAspCysGlu 93  
 QY 355 GATAATCCAAACAGCAAGAGGTTTCCAAAGAGCGGCTATCGAGATGAATPAT 414  
 Db 94 AspGlnGlnThrArgSerArgGlyPheSerLysArgGlyCysGlnAspGlyAsnAsp 113  
 QY 415 TCAGAGCTTCAAGGCCCTACAGAAAGAGTGGAAAGATATTCCAGAGTGGCCGTGA 474  
 Db 114 SerGluLysArgGlyProPheArgArgGlyLysArgLysSerPheArgGlyCysArgGly 133  
 QY 475 GGATTTGGTTCAGGAAGTCCAAATATATGATGACCCAGACAGATGATATGACGCACT 534  
 Db 134 GlyPheGlyLeuGlyArgProHisSerSerGluSerGlnAspGlnGlyThrGlnArgGly 153  
 QY 535 GGTGGCTTTTGTGTTAGAAACAGATTAATGAGTGCACAGATGATGATGATCTTCT 594  
 Db 154 GlyGlyLeuPheGlySerArgLysProHisAsnSerGlyLysAsnGlyAspThrTyr 173  
 QY 595 CAAAGCAGAAATGGCAGTGAAGTGAAGAGAGTGTTCCTCAAGGTTTAAAGAAAGTA 654  
 Db 174 GlnSerArgSerGlySerGly-----ArgLysGlyTyrLysGlyLeuAsnGluGluVal 191  
 QY 655 ATAACAGGCTCTGAAAAGATTTCTTGAAGTCAAGACAGAGAGAGAGAAAGTATGAT 714  
 Db 192 ValThrGlySerGlyLysAsnSerTrpLysSerGluThrGluGlyLysSerSerAsp 211  
 QY 715 ACTCAAGACCAAAAGTACCTACATACCCCTCTCTCACTGAGATGAGACTCCATC 774  
 Db 212 SerGlnGlyProLysValThrTyrIleProProProProGluAspGluAspSerIle 231  
 QY 775 TTTGCACTTATCAGACAGGATTAACCTTCGACAAATACGACATATTCTTGTGGAATG 834  
 Db 232 PheAlaHisTyrgLysThrGlyIleAsnAspLysTyrgAspThrIleLeuValGluVal 251  
 QY 835 TCTGCACTATGATGACCAAGCAATTCGATTTTGAAGAAGTAACTCTCTGCAGACA 894  
 Db 252 SerGlyHisAspPheLysProHisAlaIleuThrPheGluGluAlaAsnLeuCysGlnThr 271  
 QY 895 CTGAATAACAATGCTTAAAGCTGTATTAATACTTAAGTCTCTGTGCAAAATATACGT 954  
 Db 272 LeuAsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyser 291  
 QY 955 ATTCTATCATACTTGCAGAGAGATTTGATGCTTGTGCTCAAAAGGCTTGGGAG 1014  
 Db 292 IleProIleValIleuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLys 311  
 QY 1015 ACTGGGCTTTTCTCTCAATTTTGGCTCATATGATGATGATGATGAATTAAGTCCAGT 1074  
 Db 312 ThrAlaAlaPheLeuLeuProIleuAlaHisMetMetArgAspGlyIleThrAlaSer 331  
 QY 1075 GCTTTAAAGATGTCAGAGAACAGAGATGATATTATTTAGACCAACTCGAATTTGCTC 1134  
 Db 332 ArgPheLysGluLeuGlnGluProGluLysIleIleValAlaProThrArgGluLeuIle 351

QY 1135 AACCAATTTATTTGGAGCCAGAAAATTTTCTTTGGGACCTTGTAAGAGCTGTGTT 1194  
 Db 352 AsnGlnIleTyrgLeuGluAlaArgLysPheSerPheGlyThrCysValArgAlaVal 371  
 QY 1195 ATATATGGGGAAACCCAGCTGGGACATTCATTCGACAAATAGTACAGGCTGTATATA 1254  
 Db 372 IleTyrgLysIleThrGlnPheGlyHisSerValArgGlnIleValGlnGlyCysAsnIle 391  
 QY 1255 TTATGTCTACTCTCCGGAAGACATGATATCATATGCGCAAGAAAAGATGGCTCCAA 1314  
 Db 392 LeuCysAlaThrProGluArgLeuMetAspIleIleGlyLysGluLysIleGlyLeuLys 411  
 QY 1315 CAGATCAAAATACTAGTTTGGATGCAAGCTGATCGCATGTGATATGAGTTTGGTCCA 1374  
 Db 412 GlnValLysTyrgLeuValIleuAspGluAlaAspArgMetLeuAspMetGlyPheGlyPro 431  
 QY 1375 GAAATGAAGATTAATTTCTTCCAGAGATGCCATCAAGAGACAGCGCCAAACCTT 1434  
 Db 432 GluMetLysLysLeuIleSerCysProGlyMetProSerLysGluGlnArgGlnThrLeu 451  
 QY 1435 ATGTTCAAGTCACTTTTCCAGAGAAATTCAAAGGTTGGCTGCAGAGTTTAAAGTCA 1494  
 Db 452 LeuPheSerAlaThrPheProGluGluIleGlnArgLeuAlaGlyAspPheLeuLysSer 471  
 QY 1495 AATTATCTGTTGTGCTGTTTGGACAAAGTGGTGAAGATGTAGAGATGTTTCAGCAGACC 1554  
 Db 472 SerTyrgLeuPheValAlaValGlyGlnValGlyAlaCysArgAspValGlnGlnThr 491  
 QY 1555 GTTCTCCAGTGGCCAGTTCCTCAAAAGAGAAAGCTCGTTGAAATTTCTGCAAAACATA 1614  
 Db 492 IleGlnGlnValGlyGlnTyrgSerLysArgGluLysLeuValGluIleLeuArgAsnIle 511  
 QY 1615 GGGGATGAAGAACTATGCTCTTGTGGAACCTTAAGAAAACAGATTTTACTGCAACT 1674  
 Db 512 GlyAspGluArgThrMetValPheValGluThrLysLysLysAlaAspPheIleAlaThr 531  
 QY 1675 TTTCTTGTCAAGAAAAATATACATCACTACAGATCCATGATGATCGGAAACAGAGAG 1734  
 Db 532 PheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGluGlnArgGlu 551  
 QY 1735 CGGAGCAGACCTTTGGAGATTTTGGTGTGAAAGTCCAGTCTTGTGCTACTTCA 1794  
 Db 552 ArgGluGlnAlaLeuGlyLysAspPheArgCysGlyLysCysProValIleuValAlaThrSer 571  
 QY 1795 GTAGCTGCCAGAGGCTGATATTTGAATAATGTCACATGTTTCAATTTGATCTTCCCT 1854  
 Db 572 ValAlaAlaArgGlyLeuAspIleGluAsnValGlnHisValIleAsnPheAspLeuPro 591  
 QY 1855 TCTACATGATGATATGTTTCATCGAATGGGCTGACTGCTGTGCGAATATGCTGC 1914  
 Db 592 SerThrIleAspGluTyrgValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 611  
 QY 1915 AGAGCAATTTCTTTTGTGATCTTGAATCGGATTAACCATTTAGCACAGCTCTAGTAAA 1974  
 Db 612 ArgAlaIleSerPhePheAspThrAspSerAspAsnHisLeuAlaGlnProLeuValLys 631  
 QY 1975 GTATGCAAGATGTCACAGAGATGTTCTCTGCATGTTTGAAGAAAATTTGCTTATGACA 2034  
 Db 632 ValLeuSerAspAlaGlnGlnAspValProAlaTrpLeuGluGluIleAlaPheSerThr 651  
 QY 2035 TACATT---CCTGGCTCAGTGGTAGTACAAGAGAAAC---GTGTTGATCAGTTGAT 2088  
 Db 652 TyrgValProProSerPheSerSerSerThrArgGlyGlyAlaValPheAlaSerValAsp 671  
 QY 2089 ACCGAGAAG-----GGCAAGACACTTTGAACAACAGCTGGGTTTCTTTCACGA 2139  
 Db 672 ThrArgLysAsnTyrgLysLysHisThrLeuAsnThrAlaGlyIleSerSerSerGln 691  
 QY 2140 GCTTCCCATCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2172  
 Db 692 AlaProAsnProValAspAspGluSerTrpAsp 702

```

RESULT 3
DDX4_RAT ID DDX4_RAT STANDARD; PRT: 713 AA.
AC 064060;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DEAD-box protein 4 (Vasa homolog) (xVtG).
OS DDX4.
SN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistral; TISSUE=Testis;
RX MEDLINE=95160706; PubMed=7857296;
RA Komiya T., Tanigawa Y.;
RT "Cloning of a gene of the DEAD box protein family which is
RT specifically expressed in germ cells in rats.";
RL Biochem. Biophys. Res. Commun. 207:405-410(1995).
CC -1- FUNCTION: May play a role in germ cell development.
CC -1- TISSUE SPECIFICITY: Testis.
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC DDX4/VASA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S75275; AAB33364.1; -.
DR HSSP: Q58083; IHV8.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD box.
DR InterPro: IPR01650; Helicase_C.
DR Pfam: PF00270; DEAD. 1.
DR Pfam: PF00271; helicase_C; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC_C; 1.
DR PROSITE: PS00039; DEAD ATP HELICASE; 1.
KM Developmental protein: Hydrolase: ATP-binding; Repeat: Helicase.
FT NP BIND 317 324 ATP (POTENTIAL).
FT SITE 431 434 DEAD BOX.
FT DOMAIN 58 219 GLY-RICH.
SQ SEQUENCE 713 AA; 77955 MW; A31A4542EF7237F6 CRC64;

Alignment Scores:
Pred. No.: 1.29e-221 Length: 713
Score: 3247.50 Matches: 625
Percent Similarity: 90.55% Conservative: 36
Best Local Similarity: 85.62% Mismatches: 46
Query Match: 82.13% Indels: 23
DB: 1 Gaps: 6

US-09-714-865-15 (1-2172) x DDX4_RAT (1-713)
QY 1 ATGGGGGATGAAGATTGGGAAGCAATC---AACCTCATATGCTTCTCATGTTCC 57
DB 1 MetcylspergualaptpglualagluileuulysprohisvalserSerlyValPro 20
QY 58 ATATTGGAAGATAGGTAT---TCTGGAGAAAATGAGACAATTTTAAAGAGACTCCA 114
DB 21 ValpheglulysaaplystlyrSerSerGlyAlaansnglyasptnrrpheasnaArgThrSer 40
QY 115 GCTTCATCATGAGAAATGAGATGAGACCTTCGAGAGATCATTTTCATGAAAGTGA 174
DB 41 AlaaserSerSerGluMetGluaspGlyProserGlyrghasphsphenneArGserGly 60
QY 175 TTGCGCTCTGGCGGAATTTTGAAGACGAGATGCTGTGAGTGTAAATAGCGAGATAT 234

```

```

DB 61 PheSerSerGlyrghasLeuGlyAsnaArgAspIleGlyluserSerlyArGgluThr 80
QY 235 ACATCCCAATGGGTGGTTTGGAGTTGAAAAGGTTTGGAAACAGAGTTTTCAC 294
DB 81 ThrSerThrThrGlyGlyPheGlyValrGlyLyselyleGlyAsnaArgGlyPheLeuasn 100
QY 295 AGCAGGTTTGAAGATGATAGTCTGATGCTGTTCTCGAGAGAGTCTAGTAAATGACTGCCAA 354
DB 101 AsnlysheglugluglyaspserserGlyPheThrlysgluserThrmaAspCysGlu 120
QY 355 GATTAATCCAAACCGAACAAGAGGTTTCCAAAGAGCGGCTTAATGAGATGAATAT 414
DB 121 AspThrGlnThrArGserArGlyPheSerlyArGlyTyrrProaspGlyAsnaAsp 140
QY 415 TCGAAGCTTCAGGCGCATACAGAAAGGTGGAAGAGTACTGTTCCAGAGTGGCCGTGA 474
DB 141 SerGluAlaSerGlyProPheArGArgGlyGlyArGAspser----- 154
QY 475 GGATTTGCTAGAAAGTCCAAATTAATGACTTAGACCCAGACAGATATGATGACCGACT 534
DB 155 -----GluTyrrAspGlnAspGlnGlySerGlnArgGly 165
QY 535 GGTGGCTTTTGGTTCTAGAAAGACAGTATTAAGTGGACAGGTAATGTTGATCTTCT 594
DB 166 GlyGlyLeuPheGlySerArGlySproAlaAlaSerAspserGlySerGlyAspThrPhe 185
QY 595 CAAAGCAGAGTGGCAGTGGAGTGAACGAGGCGTTTCAAAAGTTTAAATGAAGAAGTA 654
DB 186 GlSerArGserGlyAsnaAla-----ArgGlyAlaTyrrlysglyLeuansnglugluVal 203
QY 655 ATAACAGGCTCTGGAAGAAATTTCTTGAGTTCAGAACGACAGAGAGAGAGAAAGTATGAT 714
DB 204 ValThrGlySerGlyLysAsnSerThrlyPheSerGluAlaGluGlyGlyluserSerAsp 223
QY 715 ACTCAAGAGCAAAAGTAGTACCTACATACCCCTCTCCACCTGAGATGAGACTCCATC 774
DB 224 IleGlnGlyProLysValThrTyrrleProProProProGluaspGluaspSerIle 243
QY 775 TTTCGACATTTATCAGACAGCATTAACCTTGACAAATACGACATATTTCTTGGAAGTG 834
DB 244 PheAlaHisTyrrGlnThrGlyIleAsnPhasPlysrThrIleleuValGluVal 263
QY 835 TCGGACATGATGACACAGCATTCGACTTGTAAAGAGTAACTCTGTCAGACA 894
DB 264 SerGlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLysCysGlnThr 283
QY 895 CTGATTAACAACATTTGCTAAAGCTGTTATTAAGTCTTACTCTGCAAAATATACAGT 954
DB 284 LeuAsnAsnAsnIleAlaLysAlaGlyTyrrThrlysheuThrProValGlnlyTyrrSer 303
QY 955 ATTCTTATCATCTTTCGACAGACAGATTTGATGGCTTGCTCAACAGGCTCTGGGAG 1014
DB 304 IleProIleValIleuAlaGlyArGAspLysMetCalaCysAlaGlnThrGlySerGlyLys 323
QY 1015 ACTGGGCTTTTCTCCATCAACAAATTTTGGCTCATATGATGATGATGAATGAATACAGCAGT 1074
DB 324 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetCetrArgAspGlyIleThrAlaSer 343
QY 1075 CGTTTAAAGAGTTGACAGAACAGAGTGTATTATTGTATACACCACTCGAGAATTTGGTC 1134
DB 344 ArgPheLysegluLeuGlnGluProGluCysIleIleValAlaProThrArGgluLeuIle 363
QY 1135 AACCAATTTATTTGGAAGCCAGAAAATTTCTTTTGGACTTGTTGTAAGCTGTTGTT 1194
DB 364 AsnGlnIleTyrrleuGluAlaArgLysPheSerPheGlyThrCysValArGAlaValVal 383
QY 1195 ATATATGGGGAAACCCAGCTGGGACATTCATTCGACCAATTAGACAGGCTGAATATA 1254
DB 384 IleTyrrGlyGlyThrGlnPheGlyHisSerIleArgGlnIleValGlnGlyCysAsnIle 403
QY 1255 TTATGTGCTACTCTCTGGAAGACTGATGATATCATAGCAAAAGAAAGATTTGCTTCAA 1314

```











## Alignment Scores:

Pred. No.:	3,84e-79	Length:	661
Score:	1229.00	Matches:	297
Percent Similarity:	57.59%	Conservative:	90
Best Local Similarity:	44.20%	Mismatches:	183
Query Match:	31.08%	Indels:	102
DB:	1	Gaps:	22

US-09-714-865-15 (1-2172) x DDX3\_MOUSE (1-661)

```

QY 214 GAGTGTAAATAGGAGATTAATACATCCACAATGGGTGTTTGAGTGGAAAGACT-- 270
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   19 AspleuanserSeraspanginSer-----GlyGlySerThrAlaSerLysGlyArg 36
QY 271 -----TTGGAACAGAGGTTTTCMAACAGACAGGTGGTGAAGTGGTGAAT 315
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   37 TyrIleProHoiHsleuArgAsnArgGlu--AlaThrLysGlyPheTyrAspLysAsp 55
QY 316 AGCTCTGGTTCTGAGAGAGTCTAGTAATGACTGCGAAGATATCCACACGGAAACAGA 375
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   56 SerSerGlyTyrPseSerSerLysAspLysAspAlaTyrSerSer----- 70
QY 376 GGGTTTCCAGAGCGCGCTATCCAGATGAAATATTCAGAGCTTGAGGCCATAC 435
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   71 ---PheGlySerArgGlyAspSerArg-----GlyLysSerSerPhePheGlyAspArg 87
QY 436 AGAAGAGGTGAAGAGTACTTTC-----CGAGGTGCGCGTGAAGATTT---GGTCTA 486
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   88 GlySerGlySerArgGlyArgPheAspAspArgGly--ArgGlyAspTyrAspGlyIle 106
QY 487 GGAAGTCCAAATATGACTTGAACCCAGACGAATGTATGACAGCGCACTGGTGCCTTTT 546
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   107 Gly----- 107
QY 547 GGTTCAGAAACAGCAGTATTAACTGCGACAGGTAAATGCTGATCTCTCAACGACAGAT 606
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   108 -----GlyArgGlyAsp-----ArgSer 113
QY 607 GGCAGTGAAGT---GAACGAGGTGTTCAAAAGTTTAAATGAAGAATATTAACAGCG 663
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   114 GlyPheGlyLysPheGlyLysArgGlyGly-----AsnSerArgTyrCysAspLys 129
QY 664 TCTGAAGAAGATCTTGAAGTCAAGAGCAGAGAGAGAGAAATAGTAGTACTCAAGA 723
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   130 SerAspGluAspAspTrpSerLys----- 137
QY 724 CCAAAAGTACCTAATACCCCTCTCTCACTGAGAGAT---GAGGACTCATCTTTGCA 780
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   138 -----ProLeuProProSerGluArgLeuGluGluGluLeuPheSer 151
QY 781 CATTATCAGACGAGCATTAACCTCGACAATAGCAACATCTTCTTGAGAGCTCGCA 840
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   152 GlyGlyAsnThrIleAsnPheGluLysTyrAspAspIleProValGluAlaThrGly 171
QY 841 CAGATGACACACAGCAATCTGACTTTGAAGAACTAATCTGTCAGACATGCAAT 900
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   172 AsnAsnCysProHoiHsIleGluSerPheSerAspValGluMetGlyGluIleMet 191
QY 901 AACCACTTGCTAAAGCTGTATATCTAAGCTTACTCTGCGAAAATAACAGATTTCT 960
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   192 GlyAsnIleGluLeuThrArgTyrThrArgProThrProValGluLysHisAlaIlePro 211
QY 961 ATATATCTGACAGAGATTGATGGCTTGTCCTCAACAGAGGTCTGGGAAGACTGCG 1020
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   212 IleIleLysGlyLysAspLeuMetAlaCysAlaGlnThrGlySerGlyLysIleAla 231
QY 1021 GCTTTCTCTACCAATTTTGCTCATATGATGATGATGCA-----ATAACT 1068
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   232 AlaPheLeuLeuProIleLeuSerGlnIleTyrAlaAspGlyProGlyGluAlaLeuArg 251
QY 1069 GCCAGTCGT-----TTTAAAGAGTTGCGAGAACCGAGTGTATTATTGTA 1113
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 252 AlaMetLysGluAsnGlyArgTyrGlyArgArgLysGlnTyrProIleSerLeuValLeu 271
QY 1114 GCACCAACTCGAAGATTGGTCAACAGATTATTATTTGAAGCCGAAATTTCTTTGGG 1173
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   272 AlaProThrArgGluLeuAlaValGlnIleTyrGluGluAlaArgLysPheSerTyrArg 291
QY 1174 ACTTGTTGAAGACTGTTGTTATATATGCGGGAACCCAGCTGGACATTCATTCACAA 1233
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   292 SerArgValArgProCysValValTyrGlyGlyAlaGluIleGlyGlnGlnIleArgAsp 311
QY 1234 ATAGTCAAGGCTGTATATATATGCTACTCTCTGGAAGACTGATGATCATAGGC 1293
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   312 LeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArgLeuValAspMetMetGlu 331
QY 1294 AAAAGAAAGATTGGTGTCAACAGATCAATCTTGTGGTGGATGAGTGTGCGATG 1353
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   332 ArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeuAspGluAlaAspArgMet 351
QY 1354 TTGATATGAGTTTGTGTCAGAAATGAAGATTATTTCTTGCCAGGAATGCCATCA 1413
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   352 LeuAspMetGlyPheGluProGlnIleArgArgIleValGluGlnAspThrMetProPro 371
QY 1414 AAGAAACAGCGCCAAACCTTATGTTCACTGCAACTTTCCAGAGAAATTCAGAGTTG 1473
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   372 LysGlyValAlaArgHisThrMetMetPheSerAlaThrPheProLysGluIleGlnMetLeu 391
QY 1474 GCTGCAGAGTTTAAAGTCAATATCTGTTGTTGCTGTTGGACAGAGGGGTGAGCA 1533
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   392 AlaArgAspPheLeu---AspGluTyrIlePheLeuAlaValGlyArgValGlySerThr 410
QY 1534 TGTAGAGATGTTCAAGACAGCCGTTCTCAAGTGGCCAGCTTCMAAAAGAGAAAGCTC 1593
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   411 SerGluAsnIleThrGlnLysValValTyrValGluGluIleAspLysArgSerPheLeu 430
QY 1594 GTTGAATTTCTCGAAACATPAGG---GATGAAGACTATGCTGTTGTTGAATTAAG 1650
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   431 LeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeuValPheValGluThrLys 450
QY 1651 AAAAAGCAGATTTCCTCCAACTTTCTTTCGAAAGAAAATATCACTCAACTATTC 1710
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   451 LysGlyAlaAspSerLeuGluAspPheLeuTyrHisGlnGlyTyrAlaCysThrSerIle 470
QY 1711 CATGTGATCGGAGACAGAGAGCGGAGAGCAAGCTCTTGAGATTTGCTTTGGTGAAG 1770
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   471 HisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHisGlnPheArgSerGlyLys 490
QY 1771 TGCCCACTTCTGTTGCTACTTCACTGAGTGCACAGGCGCTGCATTTGAATGTGCAA 1830
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   491 SerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleSerAsnValLys 510
QY 1831 CATGTATCAATTTGATCTTCTCTTACCATTTGATGATGATGATGATGATGATGATG 1890
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   511 HisValIleAsnPheAspLeuProSerAspIleGluGluTyrValHisArgIleGlyArg 530
QY 1891 ACTGTGCTGTGGGAATCTGTCAGAGCAAGCAATTTCTTTTGTGATGATCAATCGATAAC 1950
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   531 ThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePheAsn---GluArgAsnIle 549
QY 1951 CATTAGACAGGCTCTAGTAAAGATTGACAGATGCTCAACAGAGATTTCTGATGAG 2010
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   550 AsnIleThrLysAspLeuLeuAspLeuValGluAlaLysGlnGluValProSerTyrP 569
QY 2011 TTGGAAGAAATGCGCTTGTAGTACATATCTCGGCTTCAAGGTGATGTAACAGAGAAAC 2070
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   570 LeuGluAsnMetAlaPheGluHisHis-----TyrLysGlySerSerArgGlyArg 586
QY 2071 GTG-----TTGCATCAGTTGATGATGACGAAAGGCAAGAC 2106
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   587 SerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAspTyrArgGln---SerSer 605
QY 2107 ACTTGAACACAGCTGGGTTTCTTCTTCAAGACT 2142
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
   606 GlyAlaSerSerSerSerSerSerSerArgAla 617

```





```

OY 214 GAGTGTATAAGCAGATATACATCCAAATGGGTGGTTTGGAGTTGAAAGAGT--- 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 AspleuabenserserAspshnclnser-----GlyglYserThrAlaSerLysGlyArg 36
OY 271 -----TTTGAACAGAGGTTTTTCAACAGACAGGTTTGAAGATGGTAT 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 TyrilleProH1sLeuArgAsnArgGlu---AlaThrLysGlyPheTyrAspLysAsp 55
OY 316 AGCTCTGGTTTCTGAGAGAGCTAGTATATGATCGC-----GAA 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 SerSerGlyTyrSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 75
OY 355 GATATCCAAACAGACAGAGGTTTTTCCAAAGAGAGCGGCTATCCAGATGAAATATAT 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 AsperArgGlyLysSerSerPheSerAspArgGlySer----- 89
OY 415 TCAGAGAGCTTGAGGCCATACAGAAAGAGTGAAGAGTATTC-----CGAGGTTGC 468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 -----GlySerArgGlyArgPheAspAspArgGlyArg 100
OY 469 CGTGAAGATTTGGTCTAGAAAGTCCAAATATATGACTTAGCCAGACGAATGATGACAG 528
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 SerAspTyrAspGlyLysGlySer----- 108
OY 529 CGCACTGGTGCCCTTTTGGTTCTAGAAAGACCAAGTATTAGTGACAGATATGATGAT 588
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 -----ArgGlyAsp 111
OY 589 ACTTCTCAAGACAGAGAGTGGCAGTGAAGT---GAACGAGGTGGTTCAAGAGTTTAAAT 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 -----ArgSerGlyPheGlyLysPheGluArgGlyGly-----Asn 123
OY 646 GAAGAAGTAAATACAGGCTCTGGAAGAAGATTTCTTGAAGTACAGAAAGAGAGAGAGAA 705
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 SerArgTyrCysAspLysSerAspGluAspAspTyrSerLys----- 137
OY 706 AGTAGTGATCTGAAGAGCAAAAGTGACCTACATACACCCCTCCCTCCAGTATGAT--- 762
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 -----ProLeuProPheSerGluArgLeu 145
OY 763 GAGAGCTCCATCTTTCACATTATACAGACAGCATAACTTCGACAAATACAGACATAT 822
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 GluGlnGluLeuPheSerGlyGlyAsnThrGlyLeuPheGluLysLysAspPhe 165
OY 823 CTTGTGGAAGTGTCTGACATGATGACACCAACAGCAATTCGACTTTGAAGAAGCTAAT 882
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 ProValGluAlaThrGlyAsnAsnCysProProH1sLeuSerPheSerAspValGlu 185
OY 883 CTCTGTGACACACGAAATACCAATGCTTAAGGTGTTATCTAAGCTTACTCTGTG 942
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 MetGlyLysLeuIleMetGlyAsnLysLeuLeuThrArgTyrThrArgProThrProVal 205
OY 943 CAAATAATACAGTATCTTATCATATCTTGACAGACAGATTTGATGCTTGTCTCAACA 1002
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 GlnLysHisAlaIleProIleIleLysGluLysArgPheLeuMetAlaCysAlaGlnThr 225
OY 1003 GGGTCTGGGAAGACTGGCGGCTTTCTCTACCAATTTGGCTCATGATCATGATGAGA 1062
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 245
OY 1063 -----ATPACTGCCAGTGTG-----TTTAAAGATGTCAGAGAA 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 ProGlyLysAlaLeuArgAlaMetCysLysGluSerGlyArgGlyArgGlyGlnTyr 265
OY 1096 CCAGAGTGTATATATGACCAACTCGAGATTTGTCAACAGATTTATTTGGAAGCC 1155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 285
OY 1156 AGAAATTTTCTTTGGAGCTTGTGTAAAGAGCTGTTTATATATGGGGGAACCCAGCTG 1215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 ArgLysPheSerTyrArgSerArgValArgProCysValIleTyrGlyGlyAlaAspLe 305

```

```

OY 1216 GCACATTCATTCGACAAATAGTACAGAGCTGTATATATATATGCTACTCTCGAAGA 1275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 GlyGlnGlnIleArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 325
OY 1276 CTGATGAGTATCATAGGCAAGAAAGATTTGGTCTCAACAGATCAAAATCTTAGTTTG 1335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 LeuValAspMetMetGluArgGlyLysGlyLeuAspPheCysLysTyrLeuValLeu 345
OY 1336 GATGAAGCTGATGCGCATGTTGATGATGGGTTTTGTCGCAAAATGAAGAAATTAATTC 1395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgAlaGlyLeuVal 365
OY 1396 TGCCCAAGATGCCATCAAAAGAAAGACCGCAACCTTATGTTCAAGTCAACTTTTCCA 1455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 385
OY 1456 GAGAAATTCAAAGTTGGCTGCAAGATTTTAAAGTCAAAATATATGCTTTGCTGCTT 1515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 LysGlnIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePheLeuAlaVal 404
OY 1516 GCAAGTGGGTGAGAGATGATAGATGTTACAGACAGCCGTTCTCCAGATGGCCAGTTG 1575
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValAlaIleTyrValGluGluSer 424
OY 1576 TCAAAAAGAAAGAGCTCGTTGAATTCGCGAAACATAGG---GATGAAGAACTATG 1632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 AspLysArgSerPheLeuLeuAspLeuLeuAlaThrGlyLysAspSerLeuThrLeu 444
OY 1633 GTCTTTGTTGAATCTAAGAAAGAAAGAGATTTTACTGCAACTTTCTTTTCAAGAAAA 1692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 464
OY 1693 ATATCAACTACATGTATCCATGCGATCGGGAACAGAGACAGGCGGAGCAAGCTTTGGA 1752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 TyrAlaCysThrSerIleHisGlyAspArgSerGlnThrAspArgGluGluAlaLeuHis 484
OY 1753 GATTTTCGCTTTGAAAGTGCCAGTCTTGTGTTGTTACTTCACTAGTGCAGAGGCTG 1812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 GlnPheArgSerGlyLysSerProIleLeuValAlaThrValAlaAlaArgGlyLeu 504
OY 1813 GATATGAAATGTGCAACATGTTATCATATTTGATCTCTTCTTCAATGATGAATAT 1872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 524
OY 1873 GTTCATGGAATTTGGGGTACTGCTGCTGTGGGAATCACTGACAGACGAATTCCTTTT 1932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 525 ValHisArgIleGlyArgGlnArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 544
OY 1933 GATCTTGAATCGGATACCATTTAGCAGACGCTCTGATTAAGTATTTGACAGATGTCAA 1992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuValGluAlaLys 563
OY 1993 CAGAGTCTCTCGATGTTTGAAGAAATTTGCTTTAGTACATACATCTCTGCTTCAGT 2052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 GlnGluValProSerTyrLeuGluAsnMetAlaTyrGlnHisHis-----TyrLys 580
OY 2053 GGTAGTACAGAGAAACGTG-----TTTGCATCAGTGTAT 2088
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 581 GlySerSerArgLysArgSerLysSerSerArgPheSerGlyLysPheGlyAlaArgAsp 600
OY 2089 ACCAGAAAGGCAAGCACTTTGAACACAGACGCTGGGTTTTCTTCTTCACAGCT 2142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 TyrArgGln---SerSerGlyAlaSerSerSerPheSerSerSerArgAla 617

```

## RESULT 8

```

PL10 MOUSE STANDARD; PRT; 660 AA.
AC P163B1:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ATP-dependent RNA helicase PL10.
GN DIPASI OR PL10.

```



```

Db 383 ThrPheProLysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePhe 401
Qy 1507 GTTGTCTGTGGACAAGTGGGTGGAGCATGTGAAGATGTTCCAGCAGACCGTTCTCAAGTT 1566
Db 402 LeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnValValIleTyrVal 421
Qy 1567 GGCCAGTCTCAAAAAGAGAAAGAGTGTGAAATTCGCAAAACATAGG---GANGAA 1623
Db 422 GluGluAlaAspLysArgSerPheLeuLeuAspLeuAsnAlaThrGlyLysAspSer 441
Qy 1624 AGAAGTATGATGCTTTGTGTAACATGAAGAAAAGCAGATTCTTCTGCAACTTTCTTGT 1683
Db 442 LeuIleLeuValPheValGluThrLysGlyAlaAspSerLeuGluAspPheLeuTyr 461
Qy 1684 CAAGAAAAATATATCACTACAGATTCATGCTGTGTCGGGACACAGAGACGGGAGCA 1743
Db 462 HisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGlu 481
Qy 1744 GCTCTTGAGATTTTGGCTTGGAAAGTCCAGTCTTGTGTAATCTTCTGACTGAGTGGC 1803
Db 482 AlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAla 501
Qy 1804 AGAGGCTGATGATTTGAATAATGTCACATGTTATCAATTTGATCTTCTTACCAT 1863
Db 502 ArgGlyLeuAspIleSerAsenValLysHisValIleAsnPheAspLeuProSerAsp 521
Qy 1864 GATGATATGTTTCATCGAATGGCGTACTGCTGCTTGGGAATCTGCGACAGCAATT 1923
Db 522 GluGluTyrValHisArgIleGlyArgHisGlyValGlyAsnLeuGlyLeuAlaThr 541
Qy 1924 TCCTTTTGTGATCTGATCGATACGATTCATAGACAGCTCTAGTAAAGTATTGACA 1983
Db 542 SerPhePheAsn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuVal 560
Qy 1984 GATGCTCAACAGATGTTCTGATGCTGATGTTGGAAGAATGCTTATGATCATCATTCCT 2043
Db 561 GluAlaLysGlnGluValProSerTyrPheGluAsnMetAlaPheGluHis----- 578
Qy 2044 GGCTTCACTGTAGTACAAAGAGAAACGTG-----TTTGCATCA 2082
Db 579 ---TyrLysGlyGlySerArgGlyArgSerLysSerArgPheSerGlyGlyPheGlyAla 597
Qy 2083 GTTGATACGAAAGGCAAGAGCATTTGAACAGACAGTGGTTTCTTCTTCAAGAGCT 2142
Db 598 ArgAspTyrArgGln---SerSerGlyAlaSerSerSerPheSerSerGlyArgAla 616
Qy 2143 CCCAAT 2148
Db 617 SerAsn 618

RESULT 9
DDXX HUMAN STANDARD; PRT; 660 AA.
AC 015523;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN DBX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98022381; PubMed=9381176;
RA Lahn B.T., Page D.C.;
RT "Functional coherence of the human Y chromosome.";
RL Science 278:675-680(1997).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDXX3
CC SUBFAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF000985; AAC51832.1; -.
DR EMBL; AF000984; AAC51831.1; -.
DR HSSP; Q58083; IHV8.
DR Genew; HGNC:2699; DBX.
DR MIM; 400010; -.
DR InterPro; IPR00410; DEAD.
DR InterPro; IPR00629; DEAD-box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD_1.
DR Pfam; PF00271; Helicase_C_1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICG; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; ATP-binding; RNA-binding; DNA-binding.
FT NP_BIND 222 229 ATP (POTENTIAL).
FT SITE 345 348 DEAD BOX.
SQ SEQUENCE 660 AA; 73094 MW; 5A67DC082AE9CF6 CRC64;

Alignment Scores:
Pred. No.: 7.36e-76 Length: 660
Score: 1182.50 Matches: 278
Percent Similarity: 58.06% Conservative: 93
Best Local Similarity: 43.51% Mismatches: 167
Query Match: 29.91% Indels: 101
DB: 1 Gaps: 18

US-09-714-865-15 (1-2172) x DDXX HUMAN (1-660)
Qy 301 TTTCAGATGATGATAGCTGCTGTTCTGAGAGAGCTAGTATGACTGC----- 351
Db 52 PheHisAspLysAspSerSerGlyTyrSerCysSerLysAspLysAspAlaTyrSerSer 71
Qy 352 -----GAATATATCAACACGAGACAGAGGTTTCCAAAGAGCGGCTATCGA 402
Db 72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGluArgLysSer----- 89
Qy 403 GATGGAATATATTCAGAAAGCTTCAGGCGCCATACAGAAAGAGTGAAGAGTACTTC-- 459
Db 90 -----GlySerArgGlyArgPheAsp 96
Qy 460 ---CGAGTTGCCGTGAGGATTTGCTAGGAAGTCCAATATATGACTTAGACCCAGAC 516
Db 97 AspArgGly-----ArgSerAspTyrAsp----- 104
Qy 517 GAATGATAGCAGCGACTGTGCTTTTGTCTTAGAAGACCAAGTATTAACTGGCACA 576
Db 105 -----GlyIle 106
Qy 577 GGTATGCGTATATCTTCAAGACAGAGTGCAGTGAAGT---GAACGAGTGGTTAC 633
Db 107 GlyAsn-----ArgGluArgProGlyPheGlyArgPheGluArgSerGlyHis 122
Qy 634 AAAGTTTAATGAGACAGATTAATACAGGCTCTGGAAGAGATTTCTTGAAGTCAAGACA 693
Db 123 SerArgTyrProLysAspLys-----SerValGluAspArgTyrSerLys----- 126
Qy 694 GAAGAGAGAGAAAGTAGTACTCAAGACCAAAAGTAGCTTAATATCCCTCTCTCA 753
Db 137 -----ProLeuProPro 140
Qy 754 CTTGAGAT---GAGACTCATCTTTCACATTATACAGACAGCATTAATCTCGAACA 810
Db 141 SerGluArgLeuGluGlnGluPheSerGlyGlyAsnThrGlyIleAsnPheGluLys 160
Qy 811 TAGACACTATCTTGTGGAAGTGTCTGACATGATGACACACAGCAATTCGACTTTT 870

```

```

Db      161  TysAspAspIleProValGluAlaThrGlySerAsnGlyProPheHisIleGluAsnPhe 180
Qy      871  GAAGAAGCTATCTCTGACAGACATGAATACACAACTGCTAAAGCTGTTACTAAG 930
Db      181  SerAspIleAspMetGlyGluIleIleMetGlyAsnIleGluLeuThrArgIleThrArg 200
Qy      931  CTATCTCTGTCGAAAAATACAGTATCTATCATCTTCAGAGACAGATTTGATGCT 990
Db      201  ProThrProValGlnIleGlnHisAlaIleProIleIleLeuGlyLysArgAspLeuValAla 220
Qy      991  TGTGCTCAAAACAGGGTCTGGAGAGCTGGCTTTCTCTTCTACCAATTTGGCTCATATG 1050
Db      221  CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIle 240
Qy      1051  ATGCATGATGAATTAATCTGCACAGCTGTTTAAAGAGTTGACGAA----- 1095
Db      241  TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgGly 259
Qy      1096  -----CCAGAGTGATATTATTGTAGCACCAACTCGAGAAATGTCACACAG 1140
Db      260  ArgArgLysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGln 279
Qy      1141  ATTTATTGGAGCCAGAAAATTTCTTTGGACCTTGTAGAGCTGTTGTTATATAT 1200
Db      280  IleTyrGlnGluAlaArgLysPheSerTyrArgSerArgValArgProCysValIleTyr 299
Qy      1201  GGGGGAACCCAGCGTGGACATTCATTCGCAAAATATACAGGCTCTAATATATTTGT 1260
Db      300  GlyGlyAlaAspIleGlyGlnGlnIleArgAspLeuLysArgGlyCysHisLeuVal 319
Qy      1261  GCTACTCTCGAAGACTGATGATATCATATGACCAAAAGAAAGATGCTTCAACAGATC 1320
Db      320  AlathrProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCys 339
Qy      1321  AAATACTAGTTTGGATGAAGCTGATGCGATGTTGATATGGTTTGGTCCAGAAATG 1380
Db      340  LysTyrLeuValLeuAspGlnAlaAspArgMetLeuAspMetGlyPheGluProGlnIle 359
Qy      1381  AAGAGATTATTTCTTCCGACGAAATGCATCAAGAAACAGACGCCCAACCTTATATTC 1440
Db      360  ArgArgGlyLeuValGlnIleAspThrMetProProLysGlyAlaArgHisSerMetMetPhe 379
Qy      1441  AGTGCACACTTTCAGAGAAATTCAGAGGTGGCTCAGAGTTTAAAGTCAAAATAT 1500
Db      380  SerAlathrPheProLysGlnIleGlnMetLeuAlaArgAspPheLeu---AspGluTyr 398
Qy      1501  CTGTTGTTGCTGTGGACAAGTGGGTGAGCATGTAGAGATGTTTCAGACAGACGTTCTC 1560
Db      399  IlePheLeuAlaValGlyArgValGlySerThrSerGlnAsnIleThrGlnLysValVal 418
Qy      1561  CAAGTGGCCAGTTCTCAAAAAGAGAAAGCTCTTGAATTTCCGCAAAACATAGGG--- 1617
Db      419  TyrAlaIleAspLeuAspLysArgSerPheLeuLeuAspIleLeuGlyAlaThrGlySer 438
Qy      1618  GATGAAGAAGCTATGCTTTGTTGAAGTAAAGAAAGAGAGATTTTACTGCAACTTT 1677
Db      439  AspSerLeuThrLeuValAlaPheValGlnThrLysLysGlyAlaAspSerLeuGluAspPhe 458
Qy      1678  CTTTGTCAAGAAAAATATCAACTACAGATTCATGCTGATCGGAGAACAGAGACGG 1737
Db      459  LeuTyrHisGlnGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478
Qy      1738  GAGCAACCTCTTGAGAGTTTTCGCTTGGAAAGTCCAGTCTTGTGCTACTTCAGTA 1797
Db      479  GlnGlnIleLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaVal 498
Qy      1798  GCTGCCAGAGGGCTGATATATGAAGATGTCACATGTTATCAATTTGATCTCTCTCT 1857
Db      499  AlaAlaArgGlyLeuAspIleSerAsnValArgHisValIleAsnPheAspLeuProSer 518
Qy      1858  ACCATTGATGAATATGTTTCATCGAATTCGGCTATCGTGTGCGAATATCTGGCAGA 1917

```

```

Db      519  AspIleGlnGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeu 538
Qy      1918  GCAATTTCTCTTTTATCATCTTGATCGATACCATTTATGACAGCCTCTAGTAAAGTA 1977
Db      539  AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557
Qy      1978  TTGACAGATGCTCAACAGATGTTCTGATGAGTGGAGAAATTCCTTTACTATAC 2037
Db      558  LeuValGlnAlaLysGlnGluValProSerTyrLeuGlnAsnMetAlaTyrGlnHisHis 577
Qy      2038  ATTCCGCTTCTACT---GCTACTACAGAGAGAAAGCTGTTTCA-----TCA 2082
Db      578  TyrLysGlyGlySerArgGlyArgSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597
Qy      2083  GTTGATACGACGAAGGACAGACACTTTCATACACAGCTGCTTCTTCTTCACGA 2139
Db      598  ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615

RESULT 10
DED1_YEAST
ID DED1_YEAST STANDARD; PRT; 604 AA.
AC P06634;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ATP-dependent RNA helicase DED1.
GN DED1 OR SPP81 OR YOR204W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A364A X H79-20.3;
RX MEDLINE=9111585; PubMed=1996139;
RA Jamieson D.J., Rabe B., Pringle J., Beggs J.D.;
RT "A suppressor of a yeast splicing mutation (pp8-1) encodes a
RT putative ATP-dependent RNA helicase.";
RL Nature 349:715-717(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Hughes B., Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RX MEDLINE=86093663; PubMed=3001645;
RA Struhl K.;
RT "Nucleotide sequence and transcriptional mapping of the yeast
RT pet56-his3-ded1 gene region.";
RL Nucleic Acids Res. 13:8587-8601(1985).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97289744; PubMed=9144215;
RA de la Cruz J., Iost I., Kressler D., Jander P.;
RT "The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent
RT translation in Saccharomyces cerevisiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5201-5206(1997).
CC -1- FUNCTION: POSITIVELY INVOLVED IN THE INITIATION OF PROTEIN
CC SYNTHESIS. PROBABLE ATP-BINDING RNA HELICASE. ESSENTIAL FOR CELL
CC VIABILITY. SUPPRESSOR OF A PRP8 MUTATION. MAY PLAY A ROLE IN MRNA
CC SPLICING.
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and that statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X57278; CAA40546.1; -.

```







```

Db 100 -----GlyTyrSerAsnGlyGlyGlyTyrArgAsnAsnArgGlyPheGlyGln 115
Qy 619 GAACGAGGTGTTTAAAGGTTTAAATGAAGAACTTAACAGGCTTGGAAAGAACTTCT 678
Db 116 TrpAspGly-----GlnHisValIle-----GlyAlaArgAsnThr 128
Qy 679 TGAAGTCAGAACGAGAGGAGAAAGTAGTACTGACAGACCAAGCAAGTAC 728
Db 129 LeuLeuGlnArgGlnLeuPheGlyAlaValAlaAsp-----GlyThrLysValSer--- 145
Qy 739 ATACCCCTCTCCACCTGAGTAGAGTCCATCTTTGACATTAATACAGAGGATA 798
Db 146 -----ThrGlyIle 148
Qy 799 AACTTGACAAATACGACACTATTCTTGGAAGTGTCTGACATGACACCAACGCA 858
Db 149 AsnPheGlnLysTyrAspAspIleProValGlnValSerGlyGlyIleGluPro--- 167
Qy 859 ATTCTGACTTTGAAAGAGCTAATCTGTGACAGACATGAATAACACATTGGTAAAGCT 918
Db 168 ValAsnGlnPheThrSerProProLeuAsnSerHisLeuLeuGlnAsnIleLysLeuSer 187
Qy 919 GGTATATAGTACTTCTCTGTCGCAAAAATACAGTATTCCTATCATATCTTGACAGAGA 978
Db 188 GlyTyrThrGlnProThrProValGlnLysAsnSerIleProIleValThrSerGlyArg 207
Qy 979 GATTATGCGCTTGCTCAACAGGCTTGGGAAGCTCGCGCTTTCTTCTTCAACAAAT 1038
Db 208 AspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaGlyPheLeuPheProIle 227
Qy 1039 TTGGCTCATATGATGATGATGAATAACGCGAGCTTTTAAAG----- 1086
Db 228 LeuSerLeuAlaPheAspLysGlyProAlaIleValProValAspGlnAspAlaGlyMet 247
Qy 1087 -----TTGCAAGAACAGAGTATTAATTTATGATGACCACTCGAATG 1131
Db 248 GlyTyrArgProArgLysAlaTyrProThrThrLeuLeuLeuAlaProThrArgGlnLeu 267
Qy 1132 GTCAACAGATTAATTTGGAACCAAGAAATTTCTTTGGGACTTGTTGAAGCTGTT 1191
Db 268 ValCysGlnIleHisGlnGlnSerArgLysPheCysTyrArgSerThrValArgProCys 287
Qy 1192 GTTATATATGGGGAGCCAGCTGAGCATTCATTCAGCAAAATAGTACAGGCTGTAAT 1251
Db 288 AlaValIleArgLysAlaAspIleArgAlaGlnIleArgGlnIleAspGlnGlyCysAsp 307
Qy 1252 ATATTATGTCTACTCTCTGGAAGACTGATGATATCATAGCAAGAAAGATTGATCTC 1311
Db 308 LeuLeuSerAlaThrProGlyArgLeuValAspLeuIleAspArgGlyArgIleSerLeu 327
Qy 1312 AAACAGATCAAAATCTTAAGTTTGGATGAGCTGATGCGATGTGGATTTGGTTGGT 1371
Db 328 AlaAsnIleLysPheLeuValLeuAspGlnIleAspArgMetLeuAspMetGlyPheGln 347
Qy 1372 CCAGAAATGAAGAGTTAATTTCTTGCCAGAGATGCATCAAAAGCAAGCGCCAAAC 1431
Db 348 ProGlnIleArgHisIleValGlnGlyAlaAspMetThrSerValGlnGlnArgGlnThr 367
Qy 1432 CTATATGTCTAGTCACTTTTCCAGAGAAATTCAAAGTTGCTGCAAGTTTAAAG 1491
Db 368 LeuMetThrSerAlaThrPheProArgAspIleGlnLeuLeuAlaArgAspPheLeuLys 387
Qy 1492 TCAAAATATCTGTTTGTCTGCTTGAGCAAGTGGTGAGCATGTAGAGATGTTCAAG 1551
Db 388 ---AspTyrValPheLeuSerValGlyArgValIleSerThrSerGlnAsnIleThrGln 406
Qy 1552 ACCGTTCTCCAAAGTTGGCCAGCTTCCAAAGAGAGAAAGCTGTTGAATTCGCAAC 1611
Db 407 LysValAlaHisValGlnAspSerGlnLysArgSerTyrLeuLeuAspIleLeuHisThr 426
Qy 1612 ATAGGAGATGAAA---AGAACTATGCTCTTTGTTAAAATAGAAAAAGCAGATTTTACT 1668

```

```

Db 427 LeuProProGlnGlyLeuThrLeuIlePheValGlnThrLysArgMetAlaAspThrLeu 446
Qy 1669 GCAACTTTCTTTGTCACAGAAATAATATCACTACAGATTCATGATGTCGGGAACAG 1728
Db 447 ThrAspTyrLeuLeuAsnSerAsnPheProAlaIleHisIleGlyAspArgThrGln 466
Qy 1729 AGAGAGCGGAGACCAAGCTCTTGAGATTTTCGCTTTGGAAGTGGCCAGCTTCTTGCT 1788
Db 467 ArgGlnArgGlnArgAlaLeuGlnLeuPheArgSerGlyArgThrSerIleMetValAla 486
Qy 1789 ACTTCAGTACGCTCCAGAGAGGCTGATATTTGAAATATGTCACACATTTTCAATTTGAT 1848
Db 487 ThrAlaValAlaSerArgGlyLeuAspIleProAsnValThrHisValIleAsnTyrAsp 506
Qy 1849 CTTCCTTACCATGATGATATGATGATTCATGCAATTTGGCGTACTGCTGTTGGCAAT 1908
Db 507 LeuProThrAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyAsn 526
Qy 1909 ACTGCGAGACAAATTTCTTTTGTGATCTTGATCGATACCATTTAGCACAGCTCTTA 1968
Db 527 ThrGlyGlnAlaValAlaPhePheAsn---ArgAsnAsnLysGlyIleAlaLysGlnLeu 545
Qy 1969 GTAAAGATTTGACAGATCTCAACAGATGTTCTCGCATGTTGGAGAAATTTGCTTT 2028
Db 546 IleGlnLeuLeuGlnGlnIleAlaAsnGlnGlyCysProSerPheLeuIleAlaMetAlaArg 565
Qy 2029 AGTACATACATTCCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2088
Db 566 GluSer-----SerPheGlyGlyAsnGlyArgGlyArgTyrSerGlyArgGly 582
Qy 2089 ACCGAAAGGCG-----AGAGCACTTTGGAACAGCT 2121
Db 583 GlyArgGlyGlyAsnAlaTyrGlyAlaArgAspPheArgArgProThrAsnSerSer 602
Qy 2122 GCGTTTCTTCT-----TCAGAGCTGCC----- 2145
Db 603 GlyTyrSerSerGlyProSerTyrSerGlyTyrGlyGlyPheGlnSerAlaGlnThrProHis 622
Qy 2146 -----AATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2169
Db 623 HisGlyAsnThrTyrAsnSerGlySerAlaGlnSerTyr 635

RESULT 12
ID DBPL_YEAST STANDARD; PRT; 617 AA.
AC P24784; P20446;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable ATP-dependent RNA helicase DBP1 (Helicase CA1).
GN DBP1 OR YPL119C OR LPH8C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NX NCBI TaxID=4932;
RN SEQUENCE FROM N.A.
RP STRAIN=DEY319;
RC MEDLINE=91312117; PubMed=1857205;
RX MEDLINE=97313271; PubMed=9169875;
RA Buesey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Barnes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Cosker F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,

```

RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,  
 RA Humcke-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,  
 RA Komp C., Kudri O., Laishkari D., Lew H., Lin A., Lin D., Louis E.J.,  
 RA Marathe R., Messenguy F., Mewes H.-W., Mitiapati S., Mostl D.,  
 RA Meller-Buer S., Nemeth A., Newnich U., Oefner P., Pearson D.,  
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,  
 RA Scherens B., Schramm S., Schroeder M., Sidu A.M., Tettelin H.,  
 RA Urestratzu L.A., Ushinsky S., Viereendeels F., Vissers S., Voss H.,  
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.,  
 RT The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI.;  
 RL Nature 387:103-105(1997).

CC [3]  
 CC SEQUENCE OF 316-500 FROM N.A.  
 CC MEDLINE=90160368; PubMed=2406722;  
 CC Chang T.-H., Arenas J., Abelson J.;  
 CC Identification of five putative yeast RNA helicase genes.;  
 CC Proc. Natl. Acad. Sci. U.S.A. 87:1571-1575(1990).  
 CC - FUNCTION: PROBABLE ATP-BINDING RNA HELICASE. SUPPRESSOR OF A  
 CC SPP81/DED1 MUTATION.  
 CC - SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
 CC SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X55993; CA39465.1; -;  
 DR EMBL; U43503; AAB68243.1; -;  
 DR PIR; S16790; S16790.  
 DR PIR; A34848; A34848.  
 DR HSSP; OS6083; 1HV8.  
 DR SGD; S0006040; DBP1.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD\_1.  
 DR Pfam; PF00271; Helicase\_C\_1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR PROSITE; PS00039; DEAD ATP HELICASE; 1.  
 DR KMW Helicase; ATP-binding; RNA-binding.  
 FT NP BIND 198 205 ATP (BY SIMILARITY).  
 FT SITE 318 321 DEAD BOX.  
 FT CONFLICT 43 44 ST -> RS (IN REF. 1).  
 FT CONFLICT 48 48 E -> K (IN REF. 1).  
 FT CONFLICT 88 88 G -> R (IN REF. 1).  
 FT CONFLICT 115 115 E -> OK (IN REF. 1).  
 FT CONFLICT 496 496 V -> I (IN REF. 3).  
 SQ SEQUENCE 617 AA; 67917 MW; 815240487628671D CRC64;

Alignment Scores:  
 Pred. No.: 1,31e-70 Length: 617  
 Score: 1108.00 Matches: 252  
 Percent Similarity: 57.43% Conservative: 92  
 Best Local Similarity: 42.07% Mismatches: 185  
 Query Match: 28.02% Indels: 70  
 DB: 1 Gaps: 14

US-09-714-865-15 (1-2172) x DBP1\_YEAST (1-617)

QY 409 AATTAATTCAGAGCTTCAGGCGCATACAGAAAGGTGAGAGTATGTTTC----- 459  
 DB 14 AAsnAsnLysGluAsnGlyGly-----GlyGlyLysSerSerSerValProPro 30  
 QY 460 -----CGAGTTCGCGTGGAGATTTGGTTAGGAAGTCCAAATATGACTTA 507  
 DB 31 HisLeuArgSerArgGly---LysProSerPheGluArgSerThrProLysGlnGlu--- 48

QY 508 GACCCAGACAGATGATGACGCGCACTGGTGGCTTTTGGTTCTAGAAACCAATATTA 567  
 DB 49 -----AspYseVal-----ThrGlyGlyAspPheArgArg----- 59  
 QY 568 AGTGGC-----ACAGTAATGTGTACTTCTTCAAGACAGAAAGTGGCACTGAA 621  
 DB 60 AlaGlyArgGlnThrGlyAsn----- 66  
 QY 622 CGAGTGGTTTACAAAGTTTAAATGAAGATTAATACAGGCTCTGGAAAGATTTCTGG 681  
 DB 67 AsnGlyGlyPhePheGlyPheSerLysGluArgAsnGlyGlyThrSerAlaSerValAsn 86  
 QY 682 AAGTCAGAACAGACAGACAGAAAGTAAGTACTCAAGACCAAAAGTACCTACAT 741  
 DB 87 ArgGlyGlySerSerSerValLysSerSerGlyAsnArgTyrValAsnGlyLysHis 106  
 QY 742 CCC---CTTCCTCCTACCTGAGATGAGACTCCATTTTGA----- 780  
 DB 107 ProGlyProLysAsnAlaLysLeuGluAlaGluLeuPheGlyValHisAspAspProasp 126  
 QY 781 CATTATCAGACAGGCAATTAAGTTCGACAAATACGACATATTCTTGGAAAGTCTGGA 840  
 DB 127 TyrHisSerSerGlyLysPheAspAsnTyrAspAsnLeuProValAspAlaSerGly 146  
 QY 841 CATGATGACACACACCAATTCGACTTTTGAAGAGCTAATCTCTGACACACTGAAT 900  
 DB 147 LysAspValProLysProLysLeuAspPheSerSerProProLeuAspGluLeuMet 166  
 QY 901 AACCAATTCCTAAGCTGTTATTAAGCTTACTCTCTGCAAAATACATATTCTT 960  
 DB 167 GluAsnHisLysLeuAlaSerPheThrLysProThrProValGlnTyrSerLeuPro 186  
 QY 961 ATCATCTTGACAGACAGAGATTTGATGGCTTGCTGCAAAAGGCTGGGAAGACTGGC 1020  
 DB 187 IleValThrLysGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysArg 206  
 QY 1021 GCTTTTCTCTCAATTTTGGCTCATATGATGATGATGATGATGATGATGATGAT 1068  
 DB 207 GlyPheLeuPheProLeuLeuThrGluLeuPheArgSerGlyProSerProValProGlu 226  
 QY 1069 -----GCCAGTCTGTTTAAAGAGTTGCAGAAACCAAGCTGATTTATTTGACACCAACT 1122  
 DB 227 LysAlaGlnSerPheTyrSerArgLysGlyTyrProSerAlaLeuValLeuAlaProThr 246  
 QY 1123 CGAGATTTGTCACACAGATTTATTTGGAAGCCAGAAATTTCTTGGAGCTTGTGA 1182  
 DB 247 ArgGluLeuAlaThrGlnLeuPheGluAlaArgLysPheThrTyrArgSerTyrVal 266  
 QY 1183 AGAGCTGTTGTTATATATAGGGGAACCCAGCTGGGACATTCATTCAGACAAATAGTACA 1242  
 DB 267 ArgProCysValValTyrGlyGlyAlaProIleGlyAsnGlnMetArgGluValAspArg 286  
 QY 1243 GAGCTGTAATATATATATGCTACTCTCTGGAAGACTGATGATATATAGCAAAAGAAAG 1302  
 DB 287 GlyCysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuGluArgLys 306  
 QY 1303 ATTGCTCTCAACAGATCAATCAATCAATGTTTGGATGAGAGCTGATGCTGATGATG 1362  
 DB 307 ValSerLeuAlaAsnLysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMet 326  
 QY 1363 GGTTTTGGTCCAGAAATGAAGATTAATTTCTTGGCCAGAAAGTCCATCAAGAAAGAG 1422  
 DB 327 GlyPheGluProGlnLeuArgHisLeuValGluGluCysAspMetProSerValGluAsn 346  
 QY 1423 CGCAAAACCTTATATGTCGCAACTTTTCCAGAGAAATTCAAAGTTGGCTGACAG 1482  
 DB 347 ArgGlnThrLeuMetPheSerAlaThrPheProValAspIleGlnHisLeuAlaArgAsp 366  
 QY 1483 TTTTAAATCAATATATCTGTTGCTGTTGCTGTTGCAAGTGGGTGAGACATGTAGAGT 1542  
 DB 367 PheLeu---AspAsnTyrIlePheLeuSerValGlyArgValGlySerThrSerGluAsn 385  
 QY 1543 GTTCAGACAGACGCTTCTCAAGTTGGCCAGTTCTCAAAAAGAGAAAGCTGTTGAATT 1602

```

Db      386 1leThrglnArglleuTyrValAspMetAspYslyssSerAlaLeuLeuAspLeu 405
QY      1603 CTGGGAACATAGGGGATGAAAGACTATGCTCTTTGTAACATAAGAAAAAGCAT 1662
Db      406 LeuSerAlaGluHisIsglyLeuThrLeuIlePheValGluThrIlyAspMetAlaAsp 425
QY      1663 TTACTGCACTTTCTTTGTCACAAATAATCACTCAAGTATCCATGCGTATCGG 1722
Db      426 GlnLeuThrAspPheLeuIleMetGlnAspPheYsaIaThrAlaIleHISGLYAspArg 445
QY      1723 GAACAGAGAGCGGAGCAAGCTCTTGAGATTTGCGCTTTGAAAGTCCCAAGTTCTT 1782
Db      446 ThrGlnIaGlnArgIuArgAlaLeuSerAlaPheYsaIaAsnValaIaAspIleLeu 465
QY      1783 GTTGCTCTTCAGTACGTGCGCAGAGGCGCTGATTTGAAATGTCGCAACTGTTATTAAT 1842
Db      466 ValAlaThrAlaValaIaAlaArgGlyLeuAspIleProAsnValaIaThrAlaIleAsn 485
QY      1843 TTGATCTTCCTTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1902
Db      486 TyrAspLeuProSerAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAla 505
QY      1903 GGGAAATCTGCGCAGAGCAATTTCTTTTTCATCGAATCGAATAACATTAGACAG 1962
Db      506 GlyAsnThrGlyValaIaThrIserPheAsnSerAsnGlnAsn--IleValIys 524
QY      1963 CCTTACATAAAAGTATTGACAGATGCTCAACAGATGTTCTGATGCTGTTGAAGAAAT 2022
Db      525 GlyLeuMetGluIleLeuAsnGlnAlaAsnGlnIuValaIaProThrPheLeuSerAspLeu 544
QY      2023 GCGTTTGTATCATCATTCCTCGGCTTCAAGTGTATGTAAGAAGAAAGTGTTCATCA 2082
Db      545 -----Ser 545
QY      2083 GTTGATACCAAGAAAGGCAAGAGCACTTTGAACACAGCTGGTTTCTTTCACGA 2139
Db      546 ArgGlnAsnSerArgGlyGlyArgThrArgGlyGlyGlyGlyPhePheAsnSerArg 564

RESULT 13
GLH1_CABEL
ID      GLH1_CABEL STANDARD; PRT; 763 AA.
AC      P34689; O9TXH4;
DT      01-FEB-1994 (Rel. 28, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      ATP-dependent RNA helicase glh-1 (Germline helicase-1).
GN      GLH-1.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderiinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=Bristol N2;
RX      MEDLINE=9402363; PubMed=8415696;
RA      Rouseell D.L., Bennett K.L.,
RT      "Gln-1, a germ-line putative RNA helicase from Caenorhabditis, has
RT      four zinc fingers."
RT      Proc. Natl. Acad. Sci. U.S.A. 90:9300-9304(1993).
RL      [2]
RP      REVISIONS TO 83-138; 275; 288 AND 398.
RA      Rouseell D.L., McCrone U.S., Smith P.A., Gruidl M.E., Bennett K.L.,
RL      Submitted (JUL-1998) to the EMBL/Genbank/DBS databases.
CC      -1- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC      -1- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.
CC      -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC      -1- SIMILARITY: CONTAINS 4 CCHC-TYPE ZINC FINGERS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```

```

CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
Db      EMBL: L19948; AAC27384.1; -
DR      PIR; A48686; A48686.
DR      HSSP; Q58083; 1HV8.
DR      InterPro; IPR001410; DEAD.
DR      InterPro; IPR000629; DEAD box.
DR      InterPro; IPR001650; Helicase_C.
DR      InterPro; IPR001878; Znf_CCHC.
DR      Pfam; PF00098; zf-CCHC; 4.
DR      Pfam; PF00270; DEAD; 1.
DR      Pfam; PF00487; helicase_C; 1.
DR      SMART; SM00490; DEXDC; 1.
DR      SMART; SM00490; HELIC; 1.
DR      SMART; SM00343; ZNF_C2HC; 4.
DR      PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR      PROSITE; PS00158; ZF_CCHC; 4.
KM      Helicase; ATP-binding; RNA-binding; zinc-finger; Repeat.
FT      DOMAIN 24 93
FT      REPEAT 24 33
FT      REPEAT 34 43
FT      REPEAT 44 53
FT      REPEAT 54 63
FT      REPEAT 64 73
FT      REPEAT 74 83
FT      REPEAT 84 93
FT      DOMAIN 207 236
FT      ZN_FING 158 175
FT      ZN_FING 183 200
FT      ZN_FING 242 259
FT      ZN_FING 262 279
FT      ZN_FING 262 385
FT      NP_BIND 385 502
FT      SITE 499 502
SQ      SEQUENCE 763 AA; 79792 MW; ADB69DE286A028D6 CRC64;

Alignment Scores:
Pred. No.: 8,99e-70 Length: 763
Score: 1096.50 Matches: 269
Percent Similarity: 48.89% Conservative: 107
Best Local Similarity: 34.98% Mismatches: 254
Query Match: 27.73% Gaps: 139
DB: 1 Gaps: 17

US-09-714-865-15 (1-2172) x GLH1_CABEL (1-763)
QY      169 AGTGAATTTGCTCTGGGCGAAT-----TTTGAACACAGATGCTGTGAG 216
Db      32 SerGlyPheGlyGlyGlyAsnGlyGlyThrGlyPheGlyGlyGlyAsnThrGlyGly 51
QY      217 TGTAAATAGCGAGATATATCATCATCAATGGGTGTTTGGAGTTGGAAGACT----- 270
Db      52 SerGlyPheGlyGlyGlyAsnThrGlyGlySerGlyPheGlyGlyGlyAsnThrGlyGly 71
QY      271 -----TTTGAACACAGAGTTTTCACACAGAGGTTTGAAGATGTGTAT----- 315
Db      72 SerGlyPheGlyGlyGlyAsnThrCysGlySerGlyPheGlyGlyGlySerThrGlyGly 91
QY      316 -----AGCTGTGTTTCTGGAGAGAGCTAGTAATGATCGGAAGAT 357
Db      92 SerProTyrGlyGlyAlaSerSerGlyPheGlyGlySerThrAlaThrSerGlyPheGly 111
QY      358 AATCCACACGAGACAGAGGTTTTCACACAGAGGCGGTATGAAATATTAATTC 417
Db      112 SerGlyGlyIleSerSerAlaPheGlyGlySerGlyGlyPhe-----GlyGlySerAla 129
QY      418 GAGCTTCAAGGCGCATACACAGAAAGAGGTGGAAGAGTAGTTCCGAGTTGCCGTGAGCA 477
Db      130 ThrGlyPheGly-----SerGlyGlyGlySerPheGlyGlyGlyAsnSerGly 145
QY      478 TTTGGTCTAGGA----- 489

```

```

Db 146 PheGlyGlyGlyGlyHisGlyGlyGlyLeuArgAsnAsnGlyPheAsnGlyGln 165
QY 490 -----AGTCAATAATAGCTTAGACCCGACGAATGT 522
Db 166 ProGlyHisArgSerSerAspCysProGluProAlaArgGlyLeuArgGluProAlaCys 185
QY 523 ATGCAG----- 528
Db 186 TyrAsnGlyGlnProGlyHisThrSerArgGlyCysThrGluArgLysProArg 205
QY 529 -----CGCACTGGTGCCCTTTGGTCTTAGAAGACCAATTAAGTCACAGATAT 582
Db 206 GluGlyArgThrGlyGlyPheGlyGlyValAclyPheGlyAsnAsnGlyGlyAsnAsp 225
QY 583 GGTGATCTTCTCAAGCAGAAAGTGCAGATGGAATGACGAGGTGTTTCAAGAGTTTA 642
Db 226 GlyPheGlyGlyAspGlyGlyPheGlyGlyGlyGlyGlyGlyGlyPrometLysCysPhe 245
QY 643 AAT----- 645
Db 246 AsnGlyGlyGlyGlyGlyHisArgSerAlaGlyCysProGluProAlaArgGlyCysPhe 265
QY 646 -----GAAGAAGTAATACAGGCTCTGGAAGAAATTTCTGGAAGTCAAGACGAA 696
Db 266 AsnGlyGlyGlyGln-----GlyHisArgSerAsnGlyCysProAsnProAlaLys 282
QY 697 GGAAGAAAGTAAGTACTACTCAAGACCAAAAGTGAATGACCTACATACCCCTCCACCT 756
Db 283 ProArgGlyGlyValGlyGlyGlyGlyProLysAlaThrValProValGluAspAsn 302
QY 757 GAGATAGAGATCCATCTTTGCACATTAATCAGACAGGCAATTAATTGCAAAATACGAC 816
Db 303 MetGluAspValPheAsnMetGlnLysIleSerGlyGlyLeuMetPheAsnLysPhePhe 322
QY 817 ACTATCTTGAGGAAGTCTGCACATGATGCACACACGAAT-----CTGACT 867
Db 323 AspAlaGluValLysLeuThrSerSerGlyLysThrValGlyLysProCysLysThr 342
QY 868 TTGAAGAAGCTAATCTCTGCAGACACTGAATAACAATGCTGAAGCTGTATATCT 927
Db 343 PheAlaGluAlaAsnLeuThrGluThrMetGlnLysAsnValAlaHisAlaGlyTyrSer 362
QY 928 AAGCTTACTCTGTGCAAAAATACAGTATTCATCACTACTGCAAGACGAGATTGATG 987
Db 363 LysThrThrProLysGlnGlnLysAlaLeuProLeuValHisGlnGlyTyrAspIleMet 382
QY 988 GCTTGCTCTCAAAAGAGGCTCTGGAAGACCTGCGCTTTCTCCCTACCAATTTGGCTCAT 1047
Db 383 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleMetThrArg 402
QY 1048 ATGATGCATGATGGAATAACTGCACGTCGTTTAAAGAGTTGACGAACACGAGTGTATT 1107
Db 403 LeuIleAspAsnAsnLeuAsnThrAlaGlyGlyGlyCysTyrProArgCysIle 422
QY 1108 ATTGACACCACTCGAAGATTGGTCAACCAAGATTATTTGAAGCCAGAAAATTTTCT 1167
Db 423 IleLeuThrProThrArgGlyLeuAlaAspGlnIleTyrAsnGlnGlyArgLysPheAla 442
QY 1168 TTGCGAGCTTGTAAGAAGCTGTTGTTATATATGGGGGAACCCGCTGGACATTCATT 1227
Db 443 TyrGlnThrMetMetGluLysProValTyrGlyLeuAlaValGlyTyrAsnLys 462
QY 1228 CGAACAATAGTACAGGCTGTATATATATATGCTCTCTCTGGAAGCTGATGATATC 1287
Db 463 GlyGlnIleGlyLysGlyAlaThrIleIleValGlyThrValGlyArgIleLysHisPhe 482
QY 1288 ATAGGCAAGAAAAGATTGCTCAACACAGTCAATCTAGTTGGATGAAGCTGAT 1347
Db 483 CysGlyGlyGlyThrIleLysLeuAspLysCysArgPhePheValLeuAspGluAlaAsp 502
QY 1348 CGCATGTTGAT-----ATGGGTTTGGTCCGAATAATGAAGAATTATTTCTGCCCGAGA 1404
Db 503 ArgMetIleAspAlaMetGlyPheGlyThrAspIleGlnThrIleValAsnTyrAspSer 522

```

```

QY 1405 ATGCCATCAAGAGACGCCCAACCCCTTATGTTCACTGCAACTTTTCCAGAGAAATT 1464
Db 523 MetProArgLysGluAsnAlaGlnThrLeuMetPheSerAlaThrPheProAspSerVal 542
QY 1465 CAAGAGTTGGCTGCAGAGTTTAAAGTCAATATATGTTGTTGCTGTGGCAAGATG 1524
Db 543 GlnGluAlaAlaIleArgAlaPheLeuArgGluAsnTyrValMetIleAlaIleAspLysIle 562
QY 1525 GGTGAGCATGTAGAGATGTTCAAGACGACGCTTCTCCAACTTGGCCGCTCTCAAAAGA 1584
Db 563 GlyAlaAlaAsnLysCysValLeuGlnPheGluArgCysGluArgSerGlyLysLys 582
QY 1585 GAAAGCTCGTTGAATTCG-----CGA 1608
Db 583 AspLysLeuLeuGlnLeuLeuGlyIleAspIleAspSerTyrThrThrGluLysSerAla 602
QY 1609 AACATAGGAGATGAAGAACTATGCTGTTGTTGAAAATAGAAAAAGACAGATTTCAT 1668
Db 603 GluValTyrThrLysLysThrMetValPheValSerGlnArgAlaMetAlaAspIleLeu 622
QY 1669 GCAACTTTCTTGTCAGAAAATAATCAACTACAGTATCCATGCTGATCGGAGACAG 1728
Db 623 AlaSerIleLeuSerSerAlaGlnValProAlaIleThrIleHisGlyAlaArgGlyGln 642
QY 1729 AGAGAGCGGAGGACGCTCTTGAGAGATTTTCGCTTGGAAAAGTCCAGTTCTTGTGCT 1788
Db 643 ArgGluArgSerGlnAlaLeuArgGlnPheArgAsnGlySerLysProValIleAla 662
QY 1789 ACTTCACTGCTGCCAGAGGGCTGATATTAAGAAATGTCACATGTTATCAATTTGAT 1848
Db 663 ThrAlaValAlaGluArgGlyLeuAspIleLysGlyValAlaPheHisValIleAsnTyrAsp 682
QY 1849 CTTCCTTCTACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1908
Db 683 MetProAspAsnIleAspAspTyrIleHisArgIleLysArgIleLysArgIleLysArg 702
QY 1909 ACTGCGACAGCAATTCCTTTTGTGATTCGATTCGATTAACATTTAGCACAGCCCTTA 1968
Db 703 SerGlyArgAlaThrSerPhe-----IleSerGluAspCysSerLeuLeuSerGluLeu 720
QY 1969 GTAAAAGTATTGACAGATGCTCAACAGATGTTCTGCAATGTTGGAAGAAATTCCTTT 2028
Db 721 ValGlyValLeuAlaAspAlaGlnGlnIleValProAspTyrMetGln----- 736
QY 2029 AGTACATCACTTCTGCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 2088
Db 737 -----GlyAlaAlaGlyLysAsnTyrGlyAlaSer----- 746
QY 2089 ACCGAAAGGGCAAGACACTTGAACACAGCTGGGTTTCTTCTTACAGAGCTCCCAAT 2148
Db 747 -----GlyPheLysSer-----ValProThr 754
QY 2149 -----CCAGTATGATGATGATGATG 2169
Db 755 GlnValProGlnAspGlnGlyTyr 763

RESULT 14
GHH3_CAEEL
ID GHH3_CAEEL STANDARD; PRT; 720 AA.
AC 001836;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP-dependent RNA helicase gln-3 (germline helicase-3).
GN Gln-3 OR B0414.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peioderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

```

```

RX      MEDLINE=20311358; PubMed=10851135;
RA      Kuznicki L.A., Smith P.A., Leung-Chiu W.M., Estevez A.O., Scott H.C.,
RA      Bennett K.L.;
RT      "Combinatorial RNA interference indicates GLH-4 can compensate for
RT      GLH-1; these two P granule components are critical for fertility in
RT      C. elegans.";
RL      Development 127:2907-2916(2000).
RN
RP      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RL      Sammons L., Wohldmann P., Rohlfing T.;
RL      Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
CC      - FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.
CC      - DEVELOPMENTAL STAGE: DURING GERMLINE POLYMERIZATION.
CC      - SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC      - SIMILARITY: CONAINS 2 CCHC-TYPE ZINC FINGERS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

**Alignment Scores:**

Pred. No.:	1,94e-62	length:	720
Score:	992.50	Matches:	248
Percent Similarity:	51.20%	Conservative:	136
Best Local Similarity:	33.07%	Mismatches:	237
Query Match:	25.10%	Indels:	129
DB:	1	Gaps:	22

US-09-714-865-15 (1-2172) x GLH3 CAEEL (1-720)

[illegible]

```

Db      58 SerGIuLeuGIuAlaIyLsProThrIleIleSerGIuAspGIuAlaIValArgSer  77
Qy      283 -----GGTTTTTAAACAGCAGGTTTGAAAGATGGTATAGCTCTGGTTCTTGAGAGAG  336
Db      78 GIuIleGIyGIyLsPheSerGIyPheAspAspIyValAspAsnValPhe-----His  95
Qy      337 TCTAGTAATGACTCTCGGAAGATATCA-----  363
Db      96 SerAsnAsnLeuHisGIySerProSerThrThrGIuLeuGIyLsProGIyIleMet  115
Qy      364 -----ACAGCAACAGCAGGTTTTCACAAAGA  390
Db      116 AsnProArgPheLeuValGIyArgSerLeuAsnSerArgSerArgAlaValThr--Arg  134
Qy      391 GCGCGCTATTCAGATGGAATATATTCAGAAAGCTTCAGAGCCA--TACAGAAAGGTGA  447
Db      135 GIySerIyAsnGIuThrSerAsnValIyGIuAsnGIuGIySerIleHisArgSerAsp  154
Qy      448 AGAGGTAGTTTCCAGAGGTGCCGTGGAGAGATTTGGTCTAGGAAGTCCAAATATAGACTTA  507
Db      155 GIuValSerThrGIuAsnCyS-----SerAlaIyAspGIuGIuArg  168
Qy      508 GACCACAGCGAATGTATATGCAGCGGACTGTGGGCTT-----TTTGGTTACAGAGA--  558
Db      169 AspArgAsp-----SerGIyGIyValSerSerIyGIyAsnIyAspSer  183
Qy      559 -----CCAGTATTAAGTGCACAGGTAAATGGTGTACTTCTCA  597
Db      184 AspGIuPheCySGIyLThrSerProIleLeuGIuAlaIySGIyPheGIy-----Ile  200
Qy      598 AGCAGAACTGGCAGTGCAGATGAACGAGGTGCTTACAAAGTTTA-----  642
Db      201 SerAsnThrCysPheAsnCyLysIySerIyGIyHisArgAlaThrGIuCySserAlaPro  220
Qy      643 AATGAAGAAGATTAATACAGCTCTCGGAAAGAATCTTGGAAGTCAGAGCAAGCAGAGAGA  702
Db      221 GIuArgGIyCySAlaAsnCySGIyAspProAsnHisArgAlaAsnGIuCySAlaSerTrp  240
Qy      703 GAAAGTAGTGTACTCTCAAGACCA---AAAGTACCTATACACCCCTCTCCACCTGAG  759
Db      241 SerIyAsnGIyValGIuGIuProThrIyValThrIyValProValIValAspIyMet  260
Qy      760 GATGAGACCTCCATCTTTGCACATTATCAGACAGCACTTAACTTGACCAATACGACACT  819
Db      261 GIuGIuValPheSerMetLeuIyIleAsnAlaGIyAspPheAspIySAspPheAsp  280
Qy      820 ATTCTTGTGGAAGTGTCTGCACATGATGCACACAGCAAT-----CTGACTTT  870
Db      281 AlaSerValGIuLeuValSerArgGIyGIuProValThrIleGIuProCySAspSerPhe  300
Qy      871 GAAGAAGCTAATCTCTGTGCAGACACTGAAATTAACAACATTCGTTAAAGCTGTATATCAAG  930
Db      301 SerAspSerAspIleProGIuInsSerMetArgArgAsnValGIuArgAlaGIyLThrArg  320
Qy      931 CTTACTCTGTGCAAAAATACAGTATTCCTATCTACTTGCAGACAGAGATTGATGCT  990
Db      321 ThrThrProIleGIuGIuThrIleuProLeuValAlaAspGIyLysAspIleLeuAla  340
Qy      991 TGTGCTCAACAGAGGTCTGGGAAGAATCGGGCTTTTCTCTTACCAATTTTGGCTCATATG  1050
Db      341 CysAlaGIuThrIySerGIyLysThrIleAlaIlePheLeuLeuProIleMetSerArgLeu  360
Qy      1051 ATGCATGATGAATTAACGCAAGTCTTTTAAAGAGTTGCGAG-----  1092
Db      361 IleLeuGIu-----LysAspLeuAsnTrpGIyAlaGIuGIyGIy  373
Qy      1093 ---GAACCAAGGTATATTTATGTAGACCAACTGCAGATTGTCAACAGATTATTTG  1149
Db      374 CysTrpProArgCySIIeIleLeuThrProThrArgGIuLeuAlaAspGIuIleTyAsn  393
Qy      1150 GAAGCCAGAAAATTTCTTTTGGGACTTGtGTAAAGAGCTGTTGTATATATAGGGGAAC  1209
Db      394 GIuGIyArgIyLsPheSerIyGIyLSerArgMetGIuIleIySProValIyGIyGIyIle  413

```

```

QY 1210 CAGCTGGACATTCATTCGCAAAATAGTACAAAGCTGTATATATATGCTACTCT 1269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 AsnValGlyTyrAsnLysSerGlnIleMetLysGlyCysThrIleIleValGlyThrIle 433
QY 1270 GGAAGACTGATGATTCATTCATTCGCAAAAGAAAGTTGGTCTCAACAGATCAATACTTA 1339
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 GlyArgValLysHisPheCysGluAspGlyValAlaIleLysLeuAspLysCysArgTyrLeu 453
QY 1330 GTTTGGATGAGCTGATTCGATGTGGAT---ATGGGTTTGGTCCAGAAATGAAGAG 1386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 ValLeuAspGluAlaAspArgMetIleAspSerMetGlyPheGlyProGluIleGluGln 473
QY 1387 TTATTTCTTCCCGAGATGCCATCAAGAGAGAGCCCAACCTTATGTTGAGTCA 1446
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 474 IleIleAsnTyrLysAsnMetProLysAsnAspLysArgGlnThrMetMetPheSerAla 493
QY 1447 ACTTTTCCAGAGAAATTCAAAGTTGGCTGCAGAGATTTTAAAGTCAATTAATCTGTTT 1506
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 ThrPheProSerSerValGlnGluAlaAlaArgLysLeuLeuArgLysAspTyrThrMet 513
QY 1507 GTTGCTGTGGACAGAGGGGCGAGCATGTAGATGTGTCAGAGACCGTTCTCCAGTT 1566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 IleThrIleAspLysIleGlyAlaAlaAsnLysCysValIleGlnGluPheGluLeuCys 533
QY 1567 GGCCAGTTCTCAAAAGAGAGAAAGCTGCTGAATTCG----- 1605
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 534 AspArgThrSerLysValAspLysLeuLysLeuLeuGlyIleAspIleAspThrTyr 553
QY 1606 -----CGAACAATGAGGAGTAAAGAACTATGCTGTTGTTGAAACTAG 1650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 ThrThrGluLysAsnSerAspValPheValLysThrIleValPheValAlaGlnGln 573
QY 1651 AAAAAGCAGATTTTACTGCAACTTTCTTGTCAAGAAAAAATCAACTCAAGATATC 1710
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 574 LysThrAlaAspThrLeuAlaSerIleMetSerAlaAlaGlnValProAlaIleThrIle 593
QY 1711 CATGCTATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1770
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 594 HisGlyAlaArgGluGlnLysGluArgSerIleAlaLeuLysLeuPheAspGlyAla 613
QY 1771 TGCCAGTTCTTGTGCTACTTCACTAGTACGTCGAGAGGCTGATTTGAAATGTGCA 1830
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 LysProValLeuIleAlaThrAlaValAlaGluArgLysLeuAspIleLysGlyValAsp 633
QY 1831 CATGCTTCAATTTGATCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1890
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 634 HisValIleAsnTyrAspMetProAsnAsnIleAspAspTyrIleHisArgIleGlyArg 653
QY 1891 ACTGCTGTTGTGGAGATACGCGAGAGCAATTTCTTTTGGATCTTGAATCGGATAC 1950
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 654 ThrGlyArgValGlyAsnSerGlyArgAlaThrSerPheIleSerLeuAlaAspVal 673
QY 1951 CATTACACAGCCTCTAGTAAAGTATGACAGATGCTCAACAGATGTTCTGATCG 2010
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 GlnIleLeuProGlnIleValArgThrLeuAlaAspAlaGlnGlnValAlaProSerTyr 693
QY 2011 TTGGAAGAAATGCTTCTTACATACATCTCTGCTTCACTGATGATGATGATGATGATGAT 2070
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 694 MetLysGluAlaAla-----GlyGlyThrSerAsnProAsn 705
QY 2071 GTGTTT---GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2097
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 LysPheGluLysSerIleAspThrGluGln 715

```

```

GN DBP2 OR SPBP8B7.16C.
OC Schizosaccharomyces pombe (Pission yeast).
OS Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91141480; PubMed=1996094;
RA Isgo R.D., Jamieson D.J., McNeill S.A., Southgate J., McPheat J.,
RA Lane D.P.,
RT "68 RNA helicase: identification of a nucleolar form and cloning of
RT related genes containing a conserved intron in yeasts."
RL Mol. Cell. Biol. 11:1326-1333(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Matvie H., Yu G., Young D.,
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream W.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Haldago J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakeert G., Aert R., Rodden J., Gymnopreze B.,
RA Welteins I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moser D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Esfer P., Zimmermann W., Medler H., Wambut R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cernutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: NOT KNOWN. ESSENTIAL PROTEIN. MAY BE A PUTATIVE RNA
CC HELICASE.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX5/DDX17
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52648; CA316873.1; -
CC EMBL: L11574; AAA35319.1; -
CC EMBL: AL032684; CAA21801.1; -
CC PIR: S14048; S14048.
CC HSP: Q58083; IHV8.
CC InterPro: IPR001410; DEAD.
CC InterPro: IPR000629; DEAD box.
CC InterPro: IPR001650; Helicase_C.
CC Pfam: PF00270; DEAD_1.
CC Pfam: PF00271; Helicase_C_1.
CC SMART: SM00487; DEXDC; 1.

```





Wed Jun 11 10:49:01 2003

us-09-714-865-15.rsp

Page 25

Job time : 92 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 14:51:45 ; Search time 61.5 Seconds  
(without alignments)  
9412.044 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 3954  
Sequence: 1 atcg99ggatgaagatcg99a.....tagatgatgagcatg99gat 2172

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xip  
-O=/cgn2.1/USFRO.spool/US09714865/runat.05062003.111757.25999/app.query.fasta.1.2311  
-DB=A.Geneseq.101002 -QMT=fastran -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09714865 @CGN 1.1 76 @runat.05062003.111757.25999 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -NAIT -DSBLOCK=100 -LONGIOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -KGAPOP=10 -XGAPEXT=0.5 -FGADPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq.101002:\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3823	96.7	724	22	Human vasa protein
2	3247.5	82.1	713	22	Rat vasa protein
3	3174	80.3	722	22	Mouse vasa protein
4	1997	50.5	700	22	Danio reio vasa pr
5	1959.5	49.6	700	22	Frog vasa protein
6	1290	32.6	661	22	Drosophila melanog
7	1274	32.2	661	22	Fruit fly vasa pro
8	1242.5	31.4	798	22	Drosophila melanog
9	1229	31.1	662	23	ABE64631
10	1215	30.7	662	23	ABE57045
11	1215	30.7	662	19	AAW81501
12	1215	30.7	662	21	AAW64483
13	1182.5	29.9	660	19	AAW81503
14	1099	27.8	612	21	AAW17433
15	1099	27.8	612	21	AAW20655
16	1095.5	27.7	646	21	AAW17458
17	892.5	22.6	421	21	AAW17434
18	892.5	22.6	421	21	AAW20656
19	888	22.5	460	21	AAW17459
20	814	20.6	399	21	AAW17435
21	814	20.6	399	21	AAW20657
22	813.5	20.6	438	21	AAW17460
23	796.5	20.1	811	22	ABE64859
24	795.5	20.1	574	22	ABE67213
25	795.5	20.1	577	22	ABE64861
26	795.5	20.1	578	22	ABE65733
27	795.5	20.1	578	22	ABE67212
28	778.5	19.7	615	22	ABE64573
29	777.5	19.7	614	22	ABE64571
30	777.5	19.7	614	22	ABE64572
31	774.5	19.6	945	22	ABE65231
32	769	19.4	703	22	ABE63551
33	761	19.2	500	22	AAE10164
34	758	19.2	344	21	AAE96482
35	709.5	17.9	648	20	AAE96356
36	708	17.9	1275	20	AAE29084
37	704	17.8	1275	21	AAE44262
38	701	17.7	135	22	AAU23109
39	696	17.6	1224	22	ABE62191
40	691.5	17.5	791	22	ABE62354
41	681.5	17.2	709	22	AAE76859
42	681.5	17.2	709	23	AAU85514
43	681	17.2	802	22	ABE71892
44	675	17.1	628	21	AAE42879
45	675	17.1	635	20	AAE73847

## ALIGNMENTS

RESULT 1  
AAE02417  
ID AAE02417 standard; Protein; 724 AA.  
AC AAE02417;  
DT 10-AUG-2001 (first entry)  
XX  
XX Human vasa protein.  
KW Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;  
KW ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;  
KW colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;  
KW skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;  
KW medulloblastoma; choriochorionoma; squamous cell carcinoma; leukemia;  
KW acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;  
KW osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
KW leiomyosarcoma; thymodysarcoma; liposarcoma; Kaposi's sarcoma;  
KW fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;

KW teratoma; mediastinal; intracranial.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Region 1..318  
 FT Region /note= "Highly antigenic"  
 FT Region 117..146  
 FT Region /label= Epitope #1  
 FT Region 695..723  
 FT Region /label= Epitope #2  
 XX MO200136445-A1.  
 XX 25-MAY-2001.  
 XX 16-NOV-2000; 2000WO-US31485.  
 XX 18-NOV-1999; 99US-0166394.  
 XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 XX Caestrillon DH;  
 XX WPI; 2001-355606/37.  
 XX N-PSDB; AAD06354.  
 XX Novel vasa polynucleotides useful in the diagnosis or treatment of  
 PT conditions characterized by aberrant expression and/or presence of  
 PT mutant forms of vasa polynucleotides or polypeptides -  
 PS Claim 18; Page 53-54; 66pp; English.  
 XX The present sequence is human vasa protein that has germ cell specific  
 CC expression and is believed to play a determinative role in gonad  
 CC development. Germ cells are specialised to produce haploid gametes in  
 CC multicellular organisms. Vasa is useful in the diagnosis or treatment  
 CC of conditions characterised by its aberrant expression and/or the  
 CC presence of its mutant forms. The conditions include cancers such as  
 CC biliary tract, brain, breast, colon, ovarian, pancreas, prostate,  
 CC colorectal, oral, liver, lung, skin, basocellular, testis, renal,  
 CC thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,  
 CC melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
 CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
 CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
 CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
 CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
 CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
 CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
 CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
 CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
 XX Sequence 724 AA:  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 724  
 Score: 3823.00 Matches: 724  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 96.69% Gaps: 0  
 US-09-714-865-15 (1-2172) x AAB02417 (1-724)  
 QY 1 ATGGGGGATGAAGATTGGGAGAGCAAAATCAACCTCATATGTTCTTCTATGTTCCATA 60  
 Db 1 MetGlaaepGluAaepTgGluAgluAaenProHismetSerSetyValProIle 20  
 QY 61 TTTGAGAGGATAGTATTTCTGGAGAGAAATGAGACATTTTACAGAGCCCGAGCTTCA 120  
 Db 21 PheGluUlyAaepAryGlySerGlyGluAaenGlyAaenPheAaenGlyThrProAlaSer 40  
 QY 121 TCATCAGAAATGATGATGACCTTCTCGAAGAGATCATTTCAATGAAAAGTGGATTGGC 180

Db 41 SerSerGluMetAaepAaepGlyProSerAryAaepAaepHismetUlySerGlyPheAla 60  
 QY 181 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGTGAGTGTATTAAGCCAGATTAATCATCC 240  
 Db 61 SerGlyAryAaenPheGlyAaenAryAaepAlaGlyGlyCysAaenUlyAryAaenThrSer 80  
 QY 241 ACAATGGGTGGTTTGGAGTTGGAGAAAGTTTGGAAAACAGAGCTTTTCAACACAGCAG 300  
 Db 81 ThrMetGlyGlyPheGlyValGlyUlySerPheGlyAaenGlyUlyPheSerAaenSerAry 100  
 QY 301 TTTGAAGATGATATAGCTTCTGTTCTTGTGAGAGAGTCTAGTAAATGACTGCCAAGTAAT 360  
 Db 101 PheGluAaepGlyAaepSerSerGlyPheTProAryGlyUlySerSerAaenAaepCysGluAaen 120  
 QY 361 CCAACACGGAACAGAGGTTTCCAGAGAGCGGCTATCATGAGATGAAATTAATTCAGAA 420  
 Db 121 ProThrAryAaenAryGlyPheSerUlyAryGlyGlyTyrAryAaepGlyAaenAaenSerGlu 140  
 QY 421 GCTTCAAGGCGCATACAGAAAGGTGGAAGAGTATGTTCCGAGGTTCCGTTGAGAGATT 480  
 Db 141 AlaSerGlyProUlyAryAaenAryGlyAryGlySerPheAryGlyCysAryGlyPhe 160  
 QY 481 GGTCTAGAGAAAGTCCAAATTAATGACTTAGACCCAGAGAAATGTATGACGCGCACTGGTGC 540  
 Db 161 GlyUlyGlySerProAaenAaepLeuAaepProAaepGlyUlySerGlyAryThrGlyGly 180  
 QY 541 CTTTGGTCTCTGAAAGACAGCATTAATTAAGTGGCAGACAGTAAAGTGTATCTTCAAGC 600  
 Db 181 LeuPheUlySerAryAaenAryProValLeuSerGlyThrGlyAaenUlyAaenThrSerGlnSer 200  
 QY 601 AGAAGTGGCAGTGGAGAGTGAACGAGGTGTACAAAGTTTAAATGAAGAAATTAACA 660  
 Db 201 ArySerGlySerGlySerGlyUlyAryGlyGlyTyrUlyGlyUlyAaenGlyUlyValIleThr 220  
 QY 661 GGCTCTGGAAAGAAATTTCTTGGAAAGTCAAGAGCAAGAGAGAGAAAGTATGATCTCAA 720  
 Db 221 GlySerGlyUlyAaenAaenThrUlySerGlyUlyAaenGlyGlyUlyUlySerAaepThrGln 240  
 QY 721 GGACCAAAATGACCTACATACACCCCTCCACCTGAGAGATGAGAGCTCATCTTGGCA 780  
 Db 241 GlyProUlyAaenUlyAaenThrUlyLeuProProProProGluAaenUlyAaenSerIlePheAla 260  
 QY 781 CATTATCAGACAGGCAATTAACCTTGCACAAATACAGACATATTTCTTGGAAAGTGTGGA 840  
 Db 261 HisTyrGlnThrGlyUlyAaenPheAaepUlySerAaenThrIleUlyValGlyUlySerGly 280  
 QY 841 CATGATGACACACAGCAATTTCTGACTTTTGAAGAAAGCTAATCTTGTCTCAGACACTGAA 900  
 Db 281 HisAaepAlaProProAlaIleLeuThrPheGlyUlyAaenUlyCysGlnThrIleUlyAaen 300  
 QY 901 AACCAACTGCTAAAGCTGTATTAAGCTTAAGCTTACTCTCTGCAAAATACAGATTCCT 960  
 Db 301 AaenAaenIleAlaUlyAaenAaenGlyTyrThrUlyAaenThrProValGluUlyArySerIlePro 320  
 QY 961 ATCATACTTCCAGAGACAGATTTGATGCTTGTCTCAAAACAGAGTCTGGAAAGACTGCG 1020  
 Db 321 IleIleUlyAaenAaenGlyAaenPheUlyAaenAaenAlaCysAlaGlnThrGlySerGlyUlyThrAla 340  
 QY 1021 GCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGAAATTAATGCTGCACTGTTT 1080  
 Db 341 AlaPheUlyAaenProIleUlyAaenIleAaenMetCysAaepUlyIleThrAlaSerAryPhe 360  
 QY 1081 AAAGAGTTGAGAGAACAGAGGTATTAATGTTAGACCAACCTGAGAAATTTGGCAACAG 1140  
 Db 361 LysGluUlyAaenGlnUlyProGlyUlyIleValAlaProThrAryGlyUlyAaenUlyAaenGln 380  
 QY 1141 ATTATTTGAGAGCCAGAAATTTCTTGTGGACTTGTGTAAGAGCTGTGTTATATAT 1200  
 Db 381 IleTyrIleUlyAaenAaenUlyAaenPheSerPheGlyThrCysValAaenAaenValIleTyr 400  
 QY 1201 GGGGGAACCCAGCTGGGACATTCATTCGACAAATATGATACAGAGCTGTAATATATATGT 1260

```

Db      401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlyCysAsnIleLeuCys 420
QY      1261 GCTACTCTGGAAGACTGATGATATCATAGGCAAGAAAAGATGGTCTCAAAAGACATC 1320
Db      421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGlnLysIleGlyLeuLysGlnIle 440
QY      1321 AATATTACTGATTGGATGAAAGCTGATCGATGTTGGATATGGGTTTGGTCCAGAAATG 1380
Db      441 LysTyrLeuValLeuAspGlnIleAspArgMetLeuAspMetGlyPheGlyProGlnMet 460
QY      1381 AAAAGATTATTTCTGGCCAGGAATGCCATCAAGGAAGAGGCAAGGCCAAACCTTATGTTT 1440
Db      461 LysLysLeuIleSerCysProGlyMetProSerLysGlnIleGlnThrLeuMetPhe 480
QY      1441 AGTGAACATTTTCCAGAGAAATTCAGAGGTTGGCTGCAGAGTTTAAAGTCAATATAT 1500
Db      481 SerAlaThrPheProGlnGlnIleGlnArgLeuAlaIleGlnPheLeuLysSerAsnTyr 500
QY      1501 CTGTTTGGTCTGTTGGACAAGTGGGTGAGCATGTAGAGATGTTCAAGACAGCCGTTCTC 1560
Db      501 LeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrValLeu 520
QY      1561 CAAGTGGCCAGTTCCTCAAAAAGAAAGAAAGCTGTTGAAATTTCTGGCAACATAGGGGAT 1620
Db      521 GlnValGlyGlnPheSerLysArgGlnLysLeuValGlnIleLeuArgAsnIleGlyAsp 540
QY      1621 GAAAGAACTAGTCTTGTGTAACATAAGAAAAAGCAGATTTTACTGCAACTTTTCTT 1680
Db      541 GluArgThrMetValPheValGlnThrLysLysValAlaAspPheThrAlaThrPheLeu 560
QY      1681 TGTCAAGAAAAAATATCAACTACATCAAGTATCCATGTGATCGGAAACAGAGAGCGGAG 1740
Db      561 CysGlnGlnLysIleSerThrThrSerIleHisGlyAspArgGlnGlnArgGln 580
QY      1741 CAAGCTCTTGGAGATTTTGGCTTGAAGAGTCCCAAGTCTTGTGTAATTCAGTACAGT 1800
Db      581 GlnAlaLeuValLysPheArgPheGlyLysCysProValLeuValAlaThrSerValAla 600
QY      1801 GCCAGAGGCGTGATATTGAAATGTGCAACATGTTATCAATTTGATCTTCTTACCC 1860
Db      601 AlaArgGlyLeuAspIleGlnValGlnHisValIleAsnPheAspLeuProSerThr 620
QY      1861 ATTGATGAATATGTCATGCAATTTGGCGCTACTGCTGTTGTGGAAATACCTGCAAGCA 1920
Db      621 IleAspGlnLysArgValHisArgIleGlyArgThrGlyArgCysGlnSerThrGlyArgAla 640
QY      1921 ATTTCCTTTTATCTTGAATGAGATTAACCATTTAGCAAGCTCTAGTAAAGATTTG 1980
Db      641 IleSerPhePheAspLeuGlnSerAspAsnHisLeuValGlnProLeuValLysValLeu 660
QY      1981 ACAGATGCTCAACAGAGATGTTCTGTCATGCTGGAAGAAATTCCTTATGATCATTCAT 2040
Db      661 ThrAspAlaGlnIleAspValProAlaThrPheGlnGlnIleAlaPheSerThrTyrIle 680
QY      2041 CTTGGCTTCAGTGGTATGATCAAGAGAAACGTGTTGATCAGTTGATACAGAAAGGC 2100
Db      681 ProGlyPheSerGlySerThrArgGlyAsnValPheIleSerValAspThrArgLysGly 700
QY      2101 AAAGACACTTGAACACAGCTGGGTTTCTTCTTACAGACTCCCAATCCAGATGATGAT 2160
Db      701 LysSerThrLeuAsnThrAlaGlyPheSerSerSerArgAlaProAsnProValAspAsp 720
QY      2161 GAGTCATGGGAT 2172
Db      721 GluSerThrAsp 724

```

```

XX      XX      Rat vasa protein.
XX      DE
XX      Rat; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
XX      KW
XX      pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
XX      KW
XX      thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
XX      KW
XX      melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
XX      KW
XX      choriocarcinoma; squamous cell carcinoma; leukemia; acute lymphocytic;
XX      KW
XX      myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
XX      KW
XX      Acquired immune deficiency syndrome; AIDS; Bowen's disease;
XX      KW
XX      leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
XX      KW
XX      fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
XX      KW
XX      teratoma; mediastinal; intracranial.
XX      OS
XX      Rattus norvegicus.
XX      PN
XX      MO200136445-A1.
XX      PD
XX      25-MAY-2001.
XX      PF
XX      16-NOV-2000; 2000MO-US31485.
XX      PR
XX      18-NOV-1999; 99US-0166394.
XX      PA
XX      (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX      PI
XX      Castrillon DH;
XX      DR
XX      MPI; 2001-355606/37.
XX      PT
XX      Novel vasa polynucleotides useful in the diagnosis or treatment of
XX      PT
XX      conditions characterized by aberrant expression and/or presence of
XX      PT
XX      mutant forms of vasa polynucleotides or polypeptides .
XX      PS
XX      Claim 4; Page 56-58; 66pp; English.
XX      CC
XX      The present sequence is rat vasa protein. Vasa is useful in the
XX      CC
XX      diagnosis or treatment of conditions characterized by its aberrant
XX      CC
XX      expression and/or the presence of its mutant forms. The conditions
XX      CC
XX      include cancers such as biliary tract, brain, breast, colon, ovarian,
XX      CC
XX      pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
XX      CC
XX      testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
XX      CC
XX      lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
XX      CC
XX      choriocarcinoma, squamous cell carcinoma, hematological neoplasms,
XX      CC
XX      acute lymphocytic and myelogenous leukemia, multiple myeloma,
XX      CC
XX      Acquired immune deficiency syndrome (AIDS) associated leukemias,
XX      CC
XX      intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
XX      CC
XX      such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
XX      CC
XX      Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
XX      CC
XX      (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
XX      CC
XX      of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX      SQ
XX      Sequence 713 AA;

```

Alignment Scores:

Pred. No.:	Length:	713
Score:	3247.50	Matches: 625
Percent Similarity:	90.55%	Conservative: 36
Best Local Similarity:	85.62%	Mismatches: 46
Query Match:	82.13%	Indels: 23
DB:	22	Gaps: 6

```

US-09-714-865-15 (1-2172) x AAE02419 (1-713)
QY      1 ATGGGGGATGAGATTTGGGAGAGCAAAATC--AACCTCATATGCTTCTTATGTTCC 57
Db      1 MetGlyAspGlnAspIleValGlnIleLeuLysProHisValSerSerTyrValPro 20
QY      58 ATATTGAGAGAGATGATAT--TCTGAGAAAAATGAGACAAATTTTAAAGAGACTCCA 114
Db      21 ValPheGlnLysAspLysTyrSerSerGlyAlaAsnGlyAspThrPheAsnAlaGlnThrSer 40
QY      115 GCTTCATCATCAGAAATGATGATGAGACCTTCTCGAAGAGATCATTTGAAAACTGGA 174

```

```

Db      41 A l a s e r s e r g l u e t g l u a s p g l y p r o s e r g l y a r g a s p h i s p h e m e t a r g s e r g l y 60
Qy      175 T T T G C C T T G G G C G A A T T T T G G A A A C A G A G A T C T G T A G C T A A T A A C C A G A A A T 234
Db      61 P h e s e r g l y a r g a n l e u g l y a s n a r g a s p l l e g l y g u s e r l y s a r g l u t h r 80
Qy      235 A C A T C C A A T G G G T T T T G G A G T T G A A A G A G T T T T G A A A C A G A G T T T T T C A A C 294
Db      81 T h e r t h r t r g l y g l y p h e g l y a r g l y l y s e g l y p h e g l y a n a r g l y p h e l e u a s n 100
Qy      295 A C A G A G T T T A A G A G T G T A G C T G T T T T G G A G A G T C T A G T A A T G A C T G C G A A 354
Db      101 A s n l y s p h e g l u g l y a s p s e r s e r g l y p h e t r p y s g l u s e r t h r a n a s p c y s g l u 120
Qy      355 G A T P A T C C A C A C G G A C A G A G G T T T T C A A G A G A G C G C T A T C A G A T G A A A A T 414
Db      121 A s p t h r g l n t h r a r g s e r a r g l y p h e s e r l y a r g l y t y r p r o a s p g l y a s n a s p 140
Qy      415 T C A G A G C T T C A G G C C A T A C A G A A G A G T G A A G A G A G T T C C G A G T T G C C G T G A 474
Db      141 S e r g l u a s e r g l y p r o p h e a r g a r g l y c l y a r g a s p s e r ----- 154
Qy      475 G G A T T T G T C T A G A A G T C C A A T M A T G A C T T A G A C C C A G A G A T G A T G C A G C G C A C T 534
Db      155 ----- g l u t y a s p g l n a s p g l n g l y s e r g l n a r g l y 165
Qy      535 G G T G C C T T T T G T T C T A G A A G C C A G T A T T A G T G C A C A G T A T G T G A T A C T T C T 594
Db      166 G l y c l y l e u p h e g l y s e r a r g l y s p r o l a l a s e r s p s e r g l y s e r g l y a s p t h r p h e 185
Qy      595 C A A A G C A A G T G C A G T G A A G T G A A G A G G T T T A C A A A G T T T A A T G A A G A G T A 654
Db      186 G l i n s e r t s e r g l y a s n a l a ----- A r g c l y a l a t y r l y s g l y l e u a n g l u g l u a l 203
Qy      655 A T A A C A G C T C T G A A A G A A T T C T T G A A G T C A G A A C A G A G A G A G A A A G T A G T A T 714
Db      204 V a l t h r g l y s e r g l y l y s a n s e r t r p l y s e r g l u a g l u g l y g l u s e r s e r a s p 223
Qy      715 A C T A A G A C C A A A A G A G A C T A C A T A C C C C T C C A C T G A G A T G A G A T C C A T C 774
Db      224 I l e g l n l y p r o l y s v a l t h r t y r l e p r o p r o p r o p r o g l u a s p g l u a s p s e r l e 243
Qy      775 T T T G C A C A T T A T C A G A C G C A T A A C T T G C A A A T A C G A C A C T A T T C T T G T G A A G T G 834
Db      244 P h e l a h i s t y r g l n t h r g l y l e a s n p h e a s p l y s t r a s p t h r l l e l e u a l g l u a l 263
Qy      835 T C T G A C A T G A T G A C C A C C A G A A T T C G A C T T T G A A G A G T A A T C T C T G T C A G A C A 894
Db      264 S e r g l y h i s a s p a l a p r o p r o l a l l e l e u t h r p h e g l u g l u a l a a n l e u c y s g l n t h r 283
Qy      895 C T G A A T A C A C A T T G C T A A A G C T G T T A C T A G C T A C T C T G T G C A A A A T A C A G T 954
Db      284 L e u a n s n a n e n l l e a l a l y s a l a g l y t y r t h r l y s l e u t h r p r o v a l g l n l y s t y s e r 303
Qy      955 A T T C C T A C A A C T T G C A G A G A G A T T G A T G G T G C T C A A C A G G T G C G G A A G 1014
Db      304 I l e p r o l l e u a l l e a l a g l y a r g a s p l e u m e t a l a c y s a l a g l n t h r g l y s e r g l y l y s 323
Qy      1015 A C T G C G G T T T C T C A C A A T T T G C T A T A T G A T G A T G A T G A A T A C T G C A G T 1074
Db      324 T h r l a l a p h e l e u l e u p r o l l e u a l a h i s m e t e t a r g a s p g l y l l e t h r l a s e r 343
Qy      1075 C G T T T A A G A G T T G C A G A A C C A G A G T A T T A T T G A G A C C A A C T C G A A T T G T C 1134
Db      344 A r g p h e y s g l u e u g l n g l u p r o g l u c y s l l e l e u a l a p r o l t r a r g s l u l e u l l e 363
Qy      1135 A A C A G A T T T A T T G A A G C C A G A A A A T T T C T T T G G A C T T G A A G A G C T G T T G T T 1194
Db      364 A s n g l n l e t y r l e u g l u a l a r g l y s p h e s e r p h e g l y n t r c y s a l a r g a l a l a l a l 383
Qy      1195 A T A T A T G G G G A A C C A G C T G G A C A T T C A A T T G C A C A A A T A G A C G C T G A A T A T A 1254

```

---

```

Db      384 I l e t y r g l y g l y t h r c l n p h e g l y h i s e r l l e a r g l n l l e u a l g l n g l y c y s a n l l e 403
Qy      1255 T T A T G T C T A C T C T G A A G A C T G A T A T A T A G C C A A A A A G A T T G T C T C A A 1314
Db      404 L e u C y s a l a t h r p r o g l y a r g l e u m e t a s p l l e l l e g l y l y s g l u l y s l l e g l y l e u l y s 423
Qy      1315 C A G A T C A A T A C T A G T T T T G A T G A G A G T G A T G C A G T T G C A T A T G G C T T T G T C C A 1374
Db      424 G l n a l l y s t y l e u a l l e u a l l e u a s p g l u a l a a s a r g m e t l e u a s p m e t c l y p h e g l y p r o 443
Qy      1375 G A A T A A A A G A T T A T T T C T T G C C A G A A A T C C A T C A A A G A A C A G G C C A A A C C C T 1434
Db      444 G l u m e t l y s l e u l l e s e r c y s p r o g l y m e t p r o s e r l y s g l u g l n a r g l n t h r l e u 463
Qy      1435 A N T G C A G C A A C T T T T C A G A G A A T T C A A G G T T G C T G C A G A G T T T T A A G T C A 1494
Db      464 L e u p h e s e r l a t h r p h e p r o g l u g l u l e g l n a r g l e u a l a g l y l u p h e l e u l y s s e r 483
Qy      1495 A A T T A T C T G T T T G T G C T G T G A C A G T G G G T G A G A C T G T A G A C A T T C A C A G A C C 1554
Db      484 A s n t y l e u p h e a l a l a l a l a l a l g l n g l y a l a c y a r g a s p a l a g l n g l n s e r 503
Qy      1555 G T T C T C A A G T T G C C A G T T C C A A A A A G A A A A G A A A A G C T G T T G A A T T C T G C A A A C A T A 1614
Db      504 I l e u g l n l a l g l y p r o v a l p h e l y s l y a r g l y s l e u a l g l u l l e u a r g a n l l e 523
Qy      1615 G G G A T G A A A G A C T A T G G T C T T G T T G A A C T A A G A A A A A G A C A G A T T T A C T G C A C T 1674
Db      524 G l y a s p g l u a r g p r o m e t a l p h e a l g l u t n t r l y s l y l a l a s p h e l l e a l t h r 543
Qy      1675 T T T C T T T G C A A A A A A A T A T C A A C A G A T T A C A G T A T C A T G T A T C G G A C A C A G A G A G 1734
Db      544 P h e l e u c y s g l n g l u l y s l l e s e r t h r t h r s e r l l e g l y a s p a r g l u g l n a r g l u 563
Qy      1735 C G G G A C A A G C T T T T G A A G A T T T C C T T T G A A A G T C C C A G T T T T G T G T A C T T C A 1794
Db      564 A r g g l u g l n a l a l e u d l y a s p h e a r g c y s g l y l y s c y s p r o v a l l e u a l a l a t h r s e r 583
Qy      1795 G T A G C T C C A G A G G G T G A T A T T G A A A T G T G C A C A C A T T A C A A T T T G A C T T C C T 1854
Db      584 V a l l a l a l a r g l y l e u a s p l l e g l u a n a l g l n h i s v a l l e a s n p h e a n l e u p r o 603
Qy      1855 T C T A C A T T G A T G A A T A T T C A T C G A A T T G G C G T A C T G T G T T G G A A T A C T G C C 1914
Db      604 S e r t h r l l a s p l u t y r a l a h i s a r g l l e g l y a r g t h r g l y a r g c y s g l y a n t h r g l y 623
Qy      1915 A G A G C A A T T C C T T T T T G A T C T T G A A T G G A T A C C A T T T A G C A C A C C T C T A G A A A A 1974
Db      624 A r g a l a l l e s e r p h e a s p t h r g l u s e r a s p a s n h l e u a l a g l n p r o l e u a l l y s 643
Qy      1975 G T A T T G A C A G A T C T C A A C A G A T G T T C C T G C A T G T T G A A G A A A T T G C C T T A G T A C A 2034
Db      644 V a l l e u s e r a s p l a g l n g l n a s p a l p r o l a t t r p l e u g l u l l e a l a p h e s e r s e r 663
Qy      2035 T A C A T T ----- C T G C T T A G T G T A G T A C A G A G A A C G T G T T T C A T C A T T A G T A C C 2091
Db      664 T y r l a p r o p r o s e r p h e s e r a s n s e r t h r a r g l y a l a l a l a l a s e r p h e a s p t h r 683
Qy      2092 A G A A G ----- G C C A A G A C A C T T T G A C A C A C G T G G T T T C T T C T T A C A G A C T 2142
Db      684 A r g l y a s n p h e g l n g l y l y s a n t h r l e u a s t h r l a g l y l l e s e r a l a g l n a l a 703
Qy      2143 C C C A A T C C A T A G A T A G T A G T A G C A T G G A T 2172
Db      704 P r o a s n p r o v a l a s p a s p g l u s e r t r p a s p 713

```

RESULT 3  
 AAE02418  
 ID AAE02418 standard; Protein; 722 AA.  
 AC AAE02418;  
 XX  
 XX 10-AUG-2001 (first entry)

XX Mouse vasa protein.  
 XX  
 XX Mouse; vasa, therapy; cancer; oral; brain; ovarian; biliary tract; lung;  
 KM pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;  
 KM thyroid; oesophagal; endometrial; gastric; skin; liver; lymphoma;  
 KM melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;  
 KM choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;  
 KM myelogenous; multiple myeloma; Paget's disease; osteosarcoma;  
 KM Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
 KM leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
 KM fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
 KM teratoma; mediastinal; intracranial.  
 XX  
 XX Mus musculus.  
 OS  
 XX WO200136445-A1.  
 PN  
 XX 25-MAY-2001.  
 PD  
 XX 16-NOV-2000; 2000WO-US31485.  
 PF  
 XX 18-NOV-1999; 99US-0166394.  
 PR  
 XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 PA  
 XX Castillon DH;  
 PI  
 XX MPI; 2001-355606/37.  
 DR  
 XX  
 XX Novel vasa polynucleotides useful in the diagnosis or treatment of  
 PT conditions characterized by aberrant expression and/or presence of  
 PT mutant forms of vasa polynucleotides or polypeptides -  
 PS  
 XX Claim 4; Page 54-56; 66pp; English.  
 CC The present sequence is mouse vasa protein. Vasa is useful in the  
 CC diagnosis or treatment of conditions characterised by its aberrant  
 CC expression and/or the presence of its mutant forms. The conditions  
 CC include cancers such as biliary tract, brain, breast, colon, ovarian,  
 CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,  
 CC testis, renal, thyroid, cervical, endometrial, oesophagal and gastric,  
 CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
 CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
 CC acute lymphocytic and myelogenous leukaemia, multiple myeloma.  
 CC Acquired immune deficiency syndrome (AIDS) associated leukemias,  
 CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
 CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
 CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
 CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
 CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
 CC  
 XX Sequence 722 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 4,39e-294 Length: 722  
 Score: 3174.00 Matches: 611  
 Percent Similarity: 90.48% Conservative: 45  
 Best Local Similarity: 84.28% Mismatches: 59  
 Query Match: 80.27% Indels: 11  
 Gaps: 7  
 US-09-714-865-15 (1-2172) x ABE02418 (1-722)  
 QY 1 ATGGGGGATGAAATTTGGAGAGCAAAATC--AACCTCATATGCTTCTTCTATGTTCC 57  
 Db 1 MeGGlYAspGlnAspTrpGlnAGlnLeuLysProHisValSerSerTYrValPro 20  
 QY 58 ATATTGAGAGAGATGGTAT--TCTGAGAAATAGAGACAATTTTACAGACCTCCA 114  
 Db 21 ValpHeGlnLysAspLysTYrSerSerGlyAlaSerGlyAspPhrPheAsnAlgThrSer 40  
 QY 115 GCTTCATCATGAGAAATGATGATGACCTTCTCGAAGAGATCATTTTCATGAAAGTGA 174

Db 41 Ala--SerSerGlnMetGlnAspGlyProSerGlyYrGAspAspPheMetAlaSerGly 59  
 QY 175 TTGGCTCTGGGGCGGATTTTGGAAACAGAGATGCTGGAGGTAAAGCAGATTAAT 234  
 Db 60 PheProSerGlnYrGserLeuGlnSerArgAspIleGlnGlnSerSerLysGlnAsn 79  
 QY 235 ACATCCACAATGGGTGGTTTGGAGTTGGAAAGATTTTGGAAACAGAGATTTTCAAC 294  
 Db 80 ThrSerThrThrcylGlyPheGlyYrGlyGlyPheGlyYrAsnArgLysPheLysAsn 99  
 QY 295 AGCAGGTTTGAAGATGTTATGCTTGGTTTGGAGAGAGTCTGTAATGATCGCAA 354  
 Db 100 AsnLysPheGlnGlnGlnLysAspSerSerGlyPheTrpLysGlnSerAsnAsnAspCysGln 119  
 QY 355 GATTAATCCAAACCGGACAGAGGTTTCCAGAGAGGGGGCTATGAGATGGAATTAAT 414  
 Db 120 AspAsnGlnThrArgSerArgLysPheSerLysArgGlyGlyCysGlnAspLysAsnAsp 139  
 QY 415 TCAGAACTTCAGGGCCATACAGAAAGAGGTGGAAGGTAGTTCCGAGGTTGCCGTGA 474  
 Db 140 SerGlnAlaSerGlyProPheArgArgGlyYrArgGlySerPheArgGlyCysArgGly 159  
 QY 475 GGATTTGTTGAGAAAGTCCAAATTAATGACTTAGACCCAGACGAATGTATGACGCCACT 534  
 Db 160 GlyPheGlyLeuGlyYrGProAsnSerGlnSerAspGlnAspGlnGlyThrGlnCysGly 179  
 QY 535 GGTGGCTTTTGGTTCTTAGAAGACAGCATTAATTAAGTGCACAGTAAAGTATCTTCT 594  
 Db 180 GlyGlyPheLeuValLeuGlnLysProAlaAlaSerAspSerGlyYrAsnLysAspThrTYr 199  
 QY 595 CAAAGCAGAAAGTGGCGAGTGAAGTGAAGGAGGTGAACAAAGTTTAATGAAGAAGTA 654  
 Db 200 GlnSerArgSerGlySerGly-----ArgGlyGlyTYrLysGlyLeuAsnGlnGlnVal 217  
 QY 655 ATPAACAGGCTTGGAAAGATTTCTTGAAGTGAAGACAGAAAGAGAGAAATGATGAT 714  
 Db 218 ValThrGlySerGlyLysAsnSerTrpLysSerGlnThrGlnGlyGlnSerSerAsp 237  
 QY 715 ACTCAAGACCAAAAGTGAACCTATACATACCCCTCTCCACCTGAGATGAGACCTCATC 774  
 Db 238 SerGlnGlyProLysValThrTYrIleProProProProGlnAspLysSerIle 257  
 QY 775 TTGACATTAATACAGACGACCTTAACTTGAACAAATAGACACATTTCTTGGAAAGT 834  
 Db 258 PheAlaHisTYrGlnThrGlyIleAsnPheAspLysTYrAspThrIleLeuValGlnVal 277  
 QY 835 TCTGACATGATGACACACACCAATTTGACTTTTGAAGAACTAATCTCTGTCCAGACA 894  
 Db 278 SerGlyHisAspAlaProProAlaIleLeuThrPheGlnGlnAlaSerLysGlnThr 297  
 QY 895 CTGAATTAACAACATTTGCTTAAAGTGGTTTACTTAAGCTTACTCTGTGCAAAAATCAGT 954  
 Db 298 LeuAsnAsnAsnIleArgLysAlaGlyTYrThrLysLeuThrProValGlnLysTYrThr 317  
 QY 955 ATTCCATATATCTTCCAGAGAGATTTGATGGCTGTGGCTCAACAGGGTGTGGGAAG 1014  
 Db 318 IleProIleValLeuAlaGlyYrAspLysPheMetAlaCysAlaGlnThrGlnSerGlyLys 337  
 QY 1015 ACTGGCGCTTTTCTCTACCAATTTTGGCTCATATGATCATGATGAATTAATCGCAGT 1074  
 Db 338 ThrAlaAlaPheLeuLeuProIleLeuAlaHisMetMetArgAspLysIleThrAlaSer 357  
 QY 1075 CGTTTAAAGAGTTGAGAGAACCAAGGTATATTGTAGCACCAACTGAGAAATGGTC 1134  
 Db 358 ArgPheLysGlnLeuGlnLysProGlnCysIleIleValAlaProThrArgGlnLeuIle 377  
 QY 1135 AACCAATTAATTTGAGAACCAAGAAATTTCTTTGGAGCTTGATAGAGCTGTGTT 1194  
 Db 378 AsnGlnIleTYrLeuGlnAlaArgLysPheSerPheGlyThrCysValIleSerValVal 397  
 QY 1195 ATATATGGGGGAACCCAGCTGGGACATTTCAATTCGCAAAATGTATCAAGCTGTATATA 1254

```

Db      11eTy-clyglythrGlnPheGlyHisSerValArgGlnIleValGlnGlyCysAsnIle 417
Qy      1255 TTATGCTCTCTCGAAGACTGATGATATCATGGCAAGAAAAGATGGTCTCAA 1314
Db      418 LeuCysAlaThrProGlyArgLeuMetAspIleIleGlySerGlyLeuValys 437
Qy      1315 CAATCAATACTTACTGTTGATGAGACGATCGCATGTTGATATGAGGTTTGGTCCA 1374
Db      438 GlnValysTyrLeuValLeuAspGlnAlaAspSerLeuAspMetGlyPheAlaPro 457
Qy      1375 GAAATGAAGAATTATTTCTTGGCCGGAATGCCATCAAGAACGCCCAACCTT 1434
Db      458 GlnIleLysIleLeuIleSerCysProGlyMetProSerLysGlnGlnHisGlnThrLeu 477
Qy      1435 ATGTCGTCGCACTTTCCAGAGAAATCAAAAGTGGCTGAGAGCTTTAAAGTCA 1494
Db      478 LeuPheSerAlaThrPheProGlnGlnIleGlnArgLeuAlaGlyAspPheLeuLysSer 497
Qy      1495 AATTATCTGTTTGTGCTGTGGACAAGTGGGAGACATGATGAGATGTTCCAGCAGACC 1554
Db      498 AsnTyrLeuPheValAlaValGlnValGlnValGlyIleAspArgAspValGlnGlnThr 517
Qy      1555 GTTCTCCAAGTGGCCAGTTCTCAAAAAGAAAAAGCTCGTTGAATTTCTGCAACATA 1614
Db      518 IleGlnIleValGlnGlnTyrGlnIleGlyLeuLysSerLeuLeuArgPheTyrGlnAsnIle 537
Qy      1615 GGGGATGAAGAAGATATGCTGTTGTTGAACCTGAAAGAAAAGAGATTTTACTGCAACT 1674
Db      538 GlyAspGlnArgThrMetValPheValGlnThrLysLysValAspPheIleAlaThr 557
Qy      1675 TTTCTTGTCAAGAAAAAATATCACTACATCAAGTATCCATGGTGCATCGGAAACAGAGAG 1734
Db      558 PheLeuLysGlnGlnLysIleSerSerThrSerIleHisGlyAspArgGlnGlnArgGln 577
Qy      1735 CGGAGCAAGCTCTTGAGATTTTGGCTTTGGAAGTGGCCAGTTCTTGTGTTACTTCA 1794
Db      578 ArgGlnGlnAlaLeuGlnLysPheArgCysGlyLysCysProValLeuValAlaThrSer 597
Qy      1795 GTAGCTCCGAGAGGCTGATATGGAAGATGGAAGATGTTATCATTTTGTGATCTTCC 1854
Db      598 ValAlaAlaArgGlyLeuAspIleGlnAsnValGlnHisValIleAsnPheAspLeuPro 617
Qy      1855 TCTACCATGATGATATGTTGATCATGCAATGGCGTACGCTGTTGGGAATACGCGC 1914
Db      618 SerThrIleAspGlnTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGly 637
Qy      1915 AGAGCAATTTCTTTTGTGATCTTGAATCGGATACCATTTAGACAGCCTTAGTAAA 1974
Db      638 ArgAlaIleSerPhePheAspThrAspSerAspAsnHisLeuAlaGlnProLeuValLys 657
Qy      1975 GTATTGACAGATGCTCAACAGAGATGTTCCGATCGTTGGAAGAAATGCTTATAGTACA 2034
Db      658 ValLeuSerAspAlaGlnGlnAspValProAlaThrPheGlnGlnIleAlaPheSerThr 677
Qy      2035 TACATTT---CCTGGCTTCAGTGTAGTACAGAGAGAAAC---GTCTTGCATCAGTTGAT 2088
Db      678 TyrValProProSerPheSerSerSerThrArgGlyValAlaValPheAlaSerValAsp 697
Qy      2089 ACCAGAAA-----GGGCAAGAGCATTGAAACAGACCTGGGTTTCTTCTTCAGC 2138
Db      698 ThrArgLysAsnTyrGlnGlnLysAlaHisValGlnTyrSerGlyAspPhePheThr 717
Qy      2139 AGCTCCCAATCCACT 2153
Db      718 SerSerGlnSerSer 722

```

```

RESULT 4
AAE02421
ID AAE02421 standard; Procein: 700 AA.
XX AAE02421;
XX
DT 10-AUG-2001 (first entry)

```

```

XX DE Danio reio vasa protein.
XX KW Vasa: therapy; cancer; oral; brain; ovarian; biliary tract; lung;
XX KW pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
XX KW thyroid; oesophagel; endometrial; gastric; skin; liver; lymphoma;
XX KW melanoma; glioblastoma; neuroblastoma; neoplasia; medulloblastoma;
XX KW choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
XX KW myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
XX KW Acquired immune deficiency syndrome; AIDS; Bowen's disease;
XX KW leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
XX KW fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
XX KW teratoma; mediastinal; intracranial.
XX OS
XX
XX Danio reio.
XX
XX WO200136445-A1.
XX
XX 25-MAY-2001.
XX
XX 16-NOV-2000; 2000MO-US31485.
XX
XX 18-NOV-1999; 99US-0166394.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Castillon DH;
XX
XX WPI; 2001-355606/37.
XX
XX
XX Novel vasa polynucleotides useful in the diagnosis or treatment of
XX PT conditions characterized by aberrant expression and/or presence of
XX PT mutant forms of vasa polynucleotides or polypeptides -
XX
XX PS Claim 4; Page 59-61; 66pp; English.
XX
XX
XX The present sequence is Danio reio vasa protein. Vasa is useful in the
XX CC diagnosis or treatment of conditions characterised by its aberrant
XX CC expression and/or the presence of its mutant forms. The conditions
XX CC include cancers such as biliary tract, brain, breast, colon, ovarian,
XX CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
XX CC testis, renal, thyroid, cervical, endometrial, oesophagel and gastric,
XX CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
XX CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
XX CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,
XX CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,
XX CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
XX CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
XX CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
XX CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
XX CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX
XX
XX Sequence 700 AA:
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,38E-181 Length: 700
XX Score: 1997.00 Matches: 423
XX Percent Similarity: 67.37% Conservative: 89
XX Best Local Similarity: 55.66% Mismatches: 148
XX Query Match: 50.51% Indels: 100
XX DB: 22 Gaps: 20
XX
XX US-09-714-865-15 (1-2172) x AAE02421 (1-700)
XX
XX Qy 10 GAAGATTGGAGACGAGAAATCAACCTCATATGTCCTCTTATGTTCCCATATTGAGAAC 69
XX Db 2 AspAspTrpGlnGlnLysPheGlnSerProValSerCys-----Ser 15
XX
XX Qy 70 GATAGCATTTCTGAGAAAATGAGACATTTTAAACAGACCTCAGCTTCATCATCAGAA 129
XX Db 16 SerGlyPheGlyGlyAlaGlyAsnAspLys-----SerAsnSerGln 29
XX
XX Qy 130 ATGAGATGATGACCTTCTCGAAGAGATCATTTTCATGAAAAGTGATTTGCTGCGCGG 189

```



```

Db      30 GlyThrGluGlySerSerTrpLys-----MetThrGlyAsp 41
QY      190 AATTGGAAACAGAGATGCTGGTAGTAAATAGCGAGATATACATCCAAATGGT 249
      42 SerPheArgGlyArg-----Gly 47
QY      250 GGTTTTGGAGTTGGAAAGAGTTTGGAAACAGAGTTTCA-----AACAGCAGGTTT 303
Db      48 GlyArgGly---GlySerArgGlyGlyArgGlyGlyPheSerGlyPheLysSerGluIle 66
QY      304 GAAAGATGGTGATAGCTCTGCTGTTTCTGGAGAGAGCTAGTAATGACTGCGAAGATATATCCA 363
      67 AspGluAsnGlySerAspGlyGlyTrp-----AsnGlyGlyGlySerArgGly 82
QY      364 ACACGGAAACAGAGAGTTTCCAGAGAGGCGGCTATGAGATGAGAAATATATTCAGAA--- 420
Db      83 ArgGlyArgGlyGlyPhe-----ArgGlyGlyPheArgSerGlySerArgAspGluAsn 100
QY      421 -----GCTTCAGGCGCATACAGAAAGAGTGGAGAGAGT 453
Db      101 AspGluAsnGlyAsnAspAspGlyTrpLysGlyGlyGlyLysSerArgGlyArgGly 120
QY      454 AGTTTC---CGAGGTGGCCGTGAGAGATT----- 480
Db      121 GlyPheGlyGlyGlyPheArgGlyGlyGlyPheArgAspGlyGlyAsnGluAspThrGlyArg 140
QY      481 ---GGCTAGGAGAGTCCAAATATATGACTTACACCCAGCAAGATGTATGACAG----- 528
Db      141 ArgGlyPheGlyArgGlyAsnAsnGluAsnGlyAsnAspGlyGlyGlyGlyArgGly 160
QY      529 ---CGCACTGGTGCGCTTTTGGTTCTAGAAAGACAGATATTAAGTGCACAGGTAATGT 585
Db      161 ArgGlyArgGlyGlyPheArgGlyGlyPheArg-----AspGlyGlyGly 175
QY      586 GATACCTTCTCAAAACGAGAGTGCAGTGCAGTGAAGTGAAGAGTGGTTACAAAGGTTTAAT 645
Db      176 AspGlySerGlyLysArg-----GlyPheGlyArgGlyGlyPheArgGlyArgAsn 192
QY      646 GAAGAAGTATATACAGGCTCTGGAAGAATTTCTTGAAAGTCAGACAGAGAGAGAGAGAA 705
Db      193 GluGluValPheSer---LysValThrThrAlaAspLysLeuAspGluGluGlySerGlu 211
QY      706 AGTAGATGATACTCAGAGCAAAAGTAGACATACATACCCCTCCCTCAGCTGAGATGAG 765
Db      212 AsnAla-----GlyProLysValValTyrValProProProProGluGluGlu 228
QY      766 GACTCATCTTTCACATTATCAGACAGGACATAAATTGCAAAATACGACATATCTT 825
Db      229 SerSerIlePheSerHisTyrAlaThrGlyIleAsnPheAspLysTyrAspAspIleLeu 248
QY      826 GTGGAAAGTCTGAGACATGATGACCCACCGACATTTGACTTTGAAAGAGTAAATCTC 885
Db      249 ValAspValSerGlySerAsnProProLysAlaIleMetThrPheGluGluAlaGlyLeu 268
QY      886 TGTCAAGACTGATATACACATTTGCTAAAGCTGATATCTAAGCTTACTCCGTGCA 945
Db      269 CysAspSerLeuSerLysAsnValSerLysSerGlyTyrValLysProThrProValGln 288
QY      946 AAATACAGATTTCTATCATCTTGCAGAGCGAGATTTGATGGCTTGTGCTCAACAGG 1005
Db      289 LysHisGlyIleProIleIleSerAlaGlyArgAspLeuMetAlaCysAlaGlnThrGly 308
QY      1006 TCTGGGAAGACTCGCGGCTTTTCTCTTCAATTTGGCTCATATGATGATGAGATA 1065
Db      309 SerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnPheMetThrAspGlyVal 328
QY      1066 ACTGCCAGTGTATTAAAGAGTTGACAGAAACAGAGTGTATTATTAGACACCACTCGA 1125
Db      329 AlaAlaSerLysPheSerGluIleGlnGluProGluAlaIleIleValAlaAlaProThrArg 348
QY      1126 GAAATGGTCAACAGATTTATTGGAAGCGAATAATTTCTTTGGAGACTGTGTAGA 1185

```

```

Db      349 GluLeuIleAsnGlnIleTyrLeuGluAlaArgLysPheAlaTyrGlyThrCysValArg 368
QY      1186 GCTGTGTTATATATGAGGGGAGACCCAGCTGGACATTTCAATTCAGACAAATAGTACAAGC 1245
Db      369 ProValValValTyrGlyGlyIleAsnThrGlyTyrThrIleArgGluValLeuLysGly 388
QY      1246 TGTATATATATATGCTACTCTCTGAGAGACTGATGATATCATATAGCAAAAGAGATT 1305
Db      389 CysAsnValLeuCysAlaThrProGluArgLeuHisAspLeuIleGlyArgGlyLysIle 408
QY      1306 GGTCTCAACACAGATCAAAATCTTATGTTTGGATGAAGCTGATGCAATGTTGGATTGGGT 1365
Db      409 GlyLeuSerLysValArgTyrLeuValIleAspGluAlaAspArgMetLeuAspMetGly 428
QY      1366 TTGTGCCAAGATTAAGAAATTAATTTCTGTGCCAGAGATGCCATCAAAAGAACGCCG 1425
Db      429 PheGluProGluMetArgLysLeuValAlaSerProGluMetProSerLysGluLysArg 448
QY      1426 CAAACCTTATGTTTCAGTGCACCTTTTCCAGAGAAATTCAAAGTTGGCTGCAGAGTTT 1485
Db      449 GlnThrLeuMetPheSerAlaThrTyrProGluAspIleGlnArgMetAlaAlaAspPhe 468
QY      1486 TTAAAGTCAAATATCTGTTGCTGCTGTTGCAAGAGTGGGTGAGACATGTATAGATGTT 1545
Db      469 LeuLysValAspTyrIlePheLeuAlaValGlyValValGlyGlyAlaCysSerAspVal 488
QY      1546 CACGACAGCGTCTCCAGAGTGGCGAGTCTCAAAAGAGAAAGCGTGGTAAATTTCTG 1605
Db      489 GluGlnThrIleValGlnValAspGlnTyrSerLysArgAspGluLeuLeuGluLeu 508
QY      1606 CGAAGATAGGGGAGGAGAAAGAACTATGCTTTTGTGGAACATTAAGAAAAAGAGATTTT 1665
Db      509 ArgAlaThrGlyAsnGluArgThrMetValPheValGlnThrLysArgSerAlaAspPhe 528
QY      1666 ACTGCAACTTTTCTTGTCAAGAAAAATATCACTCAACATTCATGCTGTGATCGGAA 1725
Db      529 IleAlaThrPheLeuCysGlnGluLysIleSerThrSerIleHisGlyAspArgGlu 548
QY      1726 CAGAGAGCGGGAGGAGAGCTCTGGAAGTTTCTGCTTGGAAAGTCCAGTCTTGT 1785
Db      549 GlnArgGluArgGluLysAlaLeuSerAspPheArgLeuGlyHisCysProValIleVal 568
QY      1786 GCTACTTCACTAGTCCAGAGGCGTGCATATTTGAAATGTGCAACATGTTATCAATTTT 1845
Db      569 AlaThrSerValAlaAlaArgGlyLeuAspIleGluGlnValGlnHisValIleAsnPhe 588
QY      1846 GATCTCTCTTTCACATGATGATGAATATGTCATGCAATTTGGCGTACTGCTGTGGG 1905
Db      589 AspMetProSerSerIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGly 608
QY      1906 AATACCTGGCAGAGCAATTTCTTTTGTGATCTTGAAATCGGATTAACCATTTAGCACAGCT 1965
Db      609 AsnThrGlyArgAlaValSerPhePheAsnProGluSerAspThrProLeuAlaArgSer 628
QY      1966 CTAGTAAAGTATTTGACAGATGCTCAACAGAGATGCTCCGATGTTTGGAAAGATTTGCC 2025
Db      629 LeuValLysValLeuSerGlyAlaGlnGlnValValProLysTrpLeuGluGluValAla 648
QY      2026 TTATGATACATACATCTCTGCTTCAAGTGT---AGTACAAAGAGAAACGTGTTGCATCA 2082
Db      649 PheSerAlaHis-----GlyThrThrGlyPheAsnProArgGlyLysValPheAlaSer 666
QY      2083 GTTGATACCGAAGAGGCAAGGCAAGCATTGGAACACAGCTGGTTTCTTTCTTCCAGAGCT 2142
Db      667 ThrAspSerArgLysGlyLysSer-----PheLysSerAspGluPro 680
QY      2143 CCC-----AATCCAGTAGATGATGATGAGATGAGAT 2172
Db      681 ProProSerGlnThrSerAlaProSerAlaAlaAlaAlaAspAspGluGluTrpGlu 700

```

RESULT 5  
 AA02420  
 ID AA02420 standard; Protein; 700 AA.

XX	AAB02420;
AC	
DT	10-AUG-2001 (first entry)
XX	
DE	Frog vasa protein.
XX	
KM	Frog; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;
KM	pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;
KM	thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;
KM	melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;
KM	choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;
KM	myelogenous; multiple myeloma; Paget's disease; osteosarcoma;
KM	Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KM	leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KW	fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KM	teratoma; mediastinal; intracranial..
OS	
XX	Xenopus laevis.
XX	
PN	WO200136445-A1.
PD	
XX	25-MAY-2001.
XX	
PX	16-NOV-2000; 2000MO-USJ1485.
PR	
XX	18-NOV-1999; 99US-0166394.
PA	
XX	(BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
PI	Castrillon DH;
DR	
XX	WPI; 2001-355606/37.
PT	
XX	Novel vasa polynucleotides useful in the diagnosis or treatment of
PT	conditions characterized by aberrant expression and/or presence of
XX	mutant forms of vasa polynucleotides or polypeptides -
XX	
PS	Claim 4; Page 58-59; 66pp; English.
XX	
CC	The present sequence is frog vasa protein. Vasa is useful in the
CC	diagnosis or treatment of conditions characterised by its aberrant
CC	expression and/or the presence of its mutant forms. The conditions
CC	include cancers such as biliary tract, brain, breast, colon, ovarian,
CC	pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,
CC	testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,
CC	lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC	choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC	acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC	Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC	intracerebral neoplasms, Bowen's disease, Paget's disease, sarcomas
CC	such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC	Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC	(eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC	of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX	
SQ	Sequence 700 AA;
XX	
Alignment Scores:	
Pred. No.:	5,29e-178 Length: 700
Score:	1959.50 Matches: 409
Percent Similarity:	68.17% Conservative: 105
Best Local Similarity:	54.24% Mismatches: 153
Query Match:	49.56% Indels: 87
DB:	22 Gaps: 18
US-09-714-865-15 (1-2172) x AAE02420 (1-700)	
Oy	7 GATGAGATTGGGACAGCAAAATCAACCTCATATGTCTTCATATGCCATATTAG 66 :::     :::       :::
Dd	2 GTTGAATTTTphsphtreutnegutnrgulvprrothryrraIProabnphe--- 20 67 AAGGATAGTATTCTTGAGAAAATGAGACAATTTTAACAGGACTCAGCTCATCATCA 126

Db	21	----	Ser	Thr	Ileu	Glu	Thr	Glu	Ala	Met	Thr	Asp	Asn	Trp	-----	Ser	Ala	Trp	Ser	Asn	37	
Qy	127	GA	AA	TG	AT	G	A	T	G	A	C	T	T	T	CC	G	A	G	A	G	186	
Db	38	Asp	Ile	Asn	Asn	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	Glu	Met	Trp	Ser	Glu	47	
Qy	187	CG	GA	AT	T	T	T	G	AA	A	C	A	G	A	T	G	C	G	A	G	246	
Db	48	Arg	Ser	Phe	Gly	Asn	Arg	-----	Gly	Ala	Trp	Ser	Glu	Pro	Ser	Asn	Phe	-----	-----	-----	66	
Qy	247	GG	TG	T	T	T	T	G	AA	A	G	A	G	T	T	T	G	AA	A	C	306	
Db	67	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	Asn	Arg	Gly	-----	-----	73	
Qy	307	GA	NG	T	GA	T	AG	T	CT	G	T	T	T	CT	G	A	G	AG	CT	TA	351	
Db	74	Arg	Gly	Thr	Arg	Gly	Phe	Gly	Thr	Asn	Arg	Asn	Phe	Asn	Trp	Ser	Ser	Glu	Arg	Asp	93	
Qy	352	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	402	
Db	94	Val	Phe	Gly	Asp	Arg	Glu	Arg	Asp	Gly	Thr	Arg	Gly	Pro	Gly	Arg	Gly	Trp	Asn	-----	113	
Qy	403	GAT	G	AA	T	A	T	T	A	T	T	A	T	T	A	T	T	A	T	T	462	
Db	114	Gly	Asn	Glu	Met	Gly	Gln	Trp	Pro	Asn	Ala	Phe	Arg	-----	Gly	Arg	Gly	Gly	Phe	Arg	131	
Qy	463	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	495	
Db	132	Asn	Glu	Asn	Glu	Gln	Arg	Arg	Gly	Phe	Gly	Glu	Arg	Gly	Gly	Phe	Arg	Ser	Glu	Asn	151	
Qy	496	AAT	A	T	A	T	A	C	T	T	A	G	C	C	C	A	G	A	T	G	A	555
Db	152	Gln	Arg	Asn	Phe	Asp	-----	-----	-----	-----	-----	-----	-----	-----	-----	Asn	Arg	Gly	Asp	Phe	164	
Qy	556	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	591	
Db	165	Gly	Glu	Gln	Glu	Asn	Arg	Pro	Arg	Ser	Trp	Gly	Arg	Gly	Gly	Phe	Asn	Asn	Ser	Asp	184	
Qy	592	TCT	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	639	
Db	185	Gly	Gly	Arg	Arg	Gly	Arg	Gly	Gly	Arg	Gly	Gly	Arg	Gln	Trp	Gly	Gly	Trp	Gly	Gly	204	
Qy	640	TTA	A	T	A	T	A	G	A	G	A	T	T	A	T	T	A	T	T	A	699	
Db	205	Arg	Asn	Gln	Glu	Val	-----	-----	-----	-----	-----	-----	-----	-----	-----	Gly	Val	Glu	Ser	Gly	220	
Qy	700	GG	A	A	T	G	A	G	A	T	G	A	C	A	G	C	A	A	A	G	759	
Db	221	Asn	Glu	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	235	
Qy	760	GAT	G	A	G	A	C	T	C	A	T	T	T	G	A	C	A	G	A	T	819	
Db	240	Gly	Glu	Asn	Phe	Asn	Ile	Phe	Arg	Gln	Trp	Gln	Ser	Gly	Ile	Asn	Phe	Asp	Trp	-----	259	
Qy	820	ATT	CT	T	G	A	G	T	G	T	G	C	A	T	G	A	G	C	A	C	879	
Db	260	Ile	Leu	Val	Asp	Val	Thr	Gly	Leu	Asp	Val	Pro	Pro	Ala	Ile	Leu	Trp	Phe	Glu	Ala	279	
Qy	880	AAT	CT	T	G	T	G	A	C	A												

```

Db      340 GlyIleThrAlaSerGlnTyrlLeuGlnLeuGlnIleProGluAlaIleIleIleAlaPro 359
Qy      1120 ACTGAGAAATTTGTCACACGATTATTTGGACCCGAAATTTCTTTTGGAGCTTGT 1179
Db      360 ThrArgIleuLeuIleAsnGlnIleTyrlLeuAspIleArgIlePheSerTyrlGlyThrCys 379
Qy      1180 GTAAGAGCTGTTGTTATATATGCGGGAACCCAGCTGGGACATTCATTCGACAAATAGTA 1239
Db      380 ValArgProValIleValIleTyrlGlyIleGlnProValIleHisAlaMetArgAspValGlu 399
Qy      1240 CAAGGCTGATATATATATGCTACTCCCGAGAGACATGATGATATCATATGCGAAAGAA 1299
Db      400 LysGlyCysAsnIleLeuCysAlaThrProGlyIleArgLeuAspIleValSerLysGlu 419
Qy      1300 AAGATTGCTCAACAGATCAATATCTAGTTTGGATGAGTGAATCGCATTTGGAT 1359
Db      420 LysIleGlyLeuSerLysLeuArgTyrlLeuValLeuAspGluAlaAspArgMetLeuAsp 439
Qy      1360 ATGGGTTTTGGTCCAGAAATGAGAACTTAATTTCTTGCCGAGAAATGCCATCAAGGAA 1419
Db      440 MetGlyPheAlaProGluIleGluLysLeuMetThrLysProGlyMetProThrLysGlu 459
Qy      1420 CAGGCCAACCCCTTATGTCACAGCACTTTCCAGAGAAATTCAAAGTTGGCTGCA 1479
Db      460 LysArgGlnThrLeuMetPheSerAlaThrTyrlProGluGluIleArgArgLeuAlaSer 479
Qy      1480 GAGTTTAAAGTCAATATCTGTTGTTGCTGTCGACAAAGTGGGAGCATGTAGA 1539
Db      480 AsnTyrlLeuLysSerGluHisLeuPheValIleValIleGlyLeuValIleGlyAlaCysSer 499
Qy      1540 GATGTCAGACAGACCGTTTCCAAAGTTGGCCAGTTCTCAAAAGAGAAAGCTGTTGAA 1599
Db      500 AspValIleGlnThrValIleuGluMetArgGluAsnGlyLysMetGluLysLeuLeuGln 519
Qy      1600 ATTTCGGAACATAGGGGAGTGAAGAAGACTATGCTCTTTGTTGAAGTGAAGAAAAGCA 1659
Db      520 IleLeuLysSerSerGluLysGluArgThrMetIlePheValAsnThrLysLysLysLys 539
Qy      1660 GATTTACTGCACTTTCTTTCACAGAAATAATATCAATACATGATTCATGATGAT 1719
Db      540 AspPheIleAlaGlyTyrlLeuCysGlnGluLysPheSerSerThrIleHisGlyAsp 559
Qy      1720 CGGGAACAGAGACCGGAGCAAGCTTTGAGATTTTCGCTTTGGAAAGTCCCACTT 1779
Db      560 ArgGluGlnTyrlGlnArgGluSerAlaLeuThrAspPheArgThrGlyLysCysThrVal 579
Qy      1780 CTGTGCTACTCTAGTACGTCGACAGAGGCTGATTTGAAATGTGCAACATGTTATC 1839
Db      580 IleValCysThrIleAlaIleAlaIleArgGlyLeuAspIleGluAsnValIleHisValIle 599
Qy      1840 AATTTGATCTTCTTCCATGATGATGATATGTTATGATTTGGGCGTACTGCTGCT 1899
Db      600 AsnTyrlAspValProLysGluValAspGluTyrlValHisArgIleGlyArgThrGlyArg 619
Qy      1900 TGTGGAAATATGCGACAGCAATTTCTTTTATGATCTTGAATCGATTAACATTTAGCA 1959
Db      620 CysGlyAsnThrGlyLysAlaThrSerPhePheAsnValGlnAspAspHisValIleAla 639
Qy      1960 CAGCTCTAGTAAAGTATGACAGATGCTCAACAGATGTTCTGCAATGTTGGAAAGAA 2019
Db      640 ArgProLeuValLysIleLeuThrAspAlaHisGlnGluValProAlaThrLeuGlnGlu 659
Qy      2020 ATTGCTTTAGTATATCATTCCTCGGCTTCAGTGTGATGATCAAGAGAAAGTGTTCGA 2079
Db      660 IleAla-----PheIleGlyHisGlyAlaLeuAsnSerPheTyrl 672
Qy      2080 TCAGTTGATACC-----AGAAAGGCGAGAGCACTTTGAACACAGCTGGGTTTTCT 2130
Db      673 AlaIleAspSerMetGlyGluGlnAlaIleGlyLysAsnIleValIleThrThrProSerPheAla 692
Qy      2131 TCTTCAGAGCTCCCAATCCAGTGAATGATGATGATGATGGAT 2172
Db      693 GlnGlu-----GluGluAlaSerTrpAsp 700

```

```

RESULT 6
ID ABB59954
AC ABB59954;
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 6654.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX MO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL04057.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX PS Disclosure; SEQ ID NO 6654; 21pp + Sequence listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB5737-AB572072).
XX
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 661 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5.07e-114 Length: 661
XX Score: 1290.00 Matches: 314
XX Percent Similarity: 56.46% Conservative: 101
XX Best Local Similarity: 42.72% Mismatches: 230
XX Query Match: 32.63% Indels: 90
XX DB: 22 Gaps: 24
XX
XX US-09-714-865-15 (1-2172) x ABB59954 (1-661)
Qy      10 GAGATTTGGAGACAGAAATCAACCCCATATGCTTCTTATGTTCCCATATTTGAG--- 66
Db      3 AspAspTrpAspAspGlu-----ProlIleValAspThr 13
Qy      67 AAGGATAGTATTTCTGAGAGAAATGAGACATTTTAAACAGACTCCAGCTTCATCATCA 126
Db      14 ArgGlyAlaArgGlyAlaAspTrpSerAspAspGluAspThrAlaLysSerPheSerGly 33
Qy      127 GAATGAT--GATGACCTTCTCGAAGAGATCATTCATTTGAAAGATGATTTGGCTCT 183
Db      34 GluAlaGluLysAspGlyValGlyGly-----SerGlyGlyGluGly 48

```





OY	1228	CGACAAAGAGTAAACAAGGCTGTATATATATATATGTGCTACTCCCGAAAGCACTAGTATCTC	1287
Db	363	GlucyllethraArgGlyCysHisValIlealatrProGlyArgLeuLeuAspPhe	382
OY	1288	ATAGCAAAAGAAAGATTGGTGTCTCAACAGATCAAAATATTAGTTTGGATGAAGCTCAT	1347
Db	383	ValAspArgThrPheIleThrPheLeuAspThrArgPheValValLeuAspGlnIleAsp	402
OY	1348	CGCATGTGGATATATGGGTTTTGGTCCAGAAATGAAGAAAGTTAATTTCTTGCCCAAGAAATG	1407
Db	403	ArgMetLeuAspMetGlyPheSerGluAspMetArgGlnIleMetThrHisValThrMet	422
OY	1408	CCATCAAAAGCAACAGCCGCAAAACCCCTTATGTCTAGTGGCAACTTTTCCAGAGAAATTCGA	1465
Db	423	-----ArgProGlnHisGlnThrLeuMetPheSerIleAlatrPheProGlnGlnIleGln	440
OY	1468	AGGTGGCTGCGAGAGTTTTTAAAGTCAATTAATCTGTTGGTGGCTGGTGGCAAGATGGGT	1522
Db	441	ArgMetAlaGlyGlnPheLeuLys--AsnTyrValSerValAlaIleGlyIleValGly	459
OY	1528	GGAGCATGTAGAATGTTTCAGACAGCCGTTCTCCAAAGTTGGCAACTTGTCAAAAAGAGAA	1587
Db	460	GlyAlaCysSerAspValIleGlnThrIleTyrGluValAsnLysTyrAlaLysArgSer	479
OY	1588	AAAGTCGTTGAATTTCTCGCAAACTAGAGGATGAAGAAACTATAGTCTTTGTTGAAC	1647
Db	480	LysLeuIleGlnIleLeuSerGlnIleAlaAsp--GlyThrIleValPheValGlnThr	498
OY	1648	AAGAAATAAGCAGATTTTACTGCAACTTTTCTTGTCAGAAAAAATATCAACTACAGT	1707
Db	499	LysArgGlyAlaAspPheLeuAlaSerPheLeuSerGlnLysGlnPheProThrThrSer	518
OY	1708	ATCCATGGTGATCGGGAAACAGAGAGCCGGACCAAGCTCTTGAGATTTTCGCTTTGGA	1767
Db	519	IleHisGlyAspArgLeuGlnSerGlnArgGlnIleAlaLeuArgAspPheLysAsnGly	538
OY	1768	AAAGTCCCAAGTTCTTTGTTGCTACTTCACTTCAAGTAGTGCACAGAGGCTGGATTTGAAAATGTG	1827
Db	539	SerMetLysValLeuIleAlaThrSerValAlaSerArgLysAspIleLysAsnIle	558
OY	1828	CAACATGTATCAATTTGATGCTTCTCTTACCATGTGAATATGTTTCATCGAATTTGGG	1887
Db	559	LysHisValIleAsnTyrAspMetProSerLysIleAspAspTyrValHisArgIleGly	578
OY	1888	CGTACTGTCGTTGGGAAATCTGCGACAGAGCAATTTCTTTTGGATCTTGAATCGGAT	1947
Db	579	ArgThrGlyCysValGlyAsnAsnGlnArgAlaThrSerPhePheAspProGlnLysAsp	598
OY	1948	AAACATTTAGCACACGCTCTAGTAAAGTATTTGACAGATGCTCAACAGAGATTTCTTGCA	2007
Db	599	ArgAlaIleAlaIleAspLeuValLysIleLeuGlnGlySerGlnThrValProAsp	618
OY	2008	TGGTTGGAAAGAAATTCCTTTAGTACATACATCTCGCTGCTCAGTGGTGAAGTCAAGAGA	2067
Db	619	PheLeuArg-----ThrCysGlyAlaGlyGlyAspGlyGlyTyrSerAsn	633
OY	2068	AAAGTGTTCATCAGTTGATATCCAGAAAAGGCAAGACACTTTAAACAAGCTGGTTT	2127
Db	634	GlnAsnPheGlyGlyValAspValArg--GlyArgGly-----AsnTyrValGly---	649
OY	2128	TCTTCTTCAAGAGCTCCCAATCCAGTACAGATGATGATCATCGGAT	2172
Db	650	-----AspAlaThrAsnValGlnGlnGlnGlnGlnIntPasp	661
RESULT 8			
ABBe4631			
ID ABBe4631 standard; Protein; 798 AA.			
XX ABBe4631;			
DT 26-MAR-2002 (first entry)			
XX Drosophila melanogaster polypeptide SEQ ID NO 20685.			

[illegible]







```

Db      412 SerGluSnnlethrglnysValtrpValGluGluLeaSpIysArgSerPheLeu 431
QY      1594 GTTGAAATTCGCAACATAGG---GATGAAAGACTATGCTGTTGAAACTAG 1650
Db      432 LeuAspLeuLeuAenAlatrGlyLysAspSerLeuThrLeuValPheValGluThrLys 451
QY      1651 AAAAAACGACATTTTACTGCAACTTTCTTTGTCAGAAAAAATATCACTACATGATC 1710
Db      452 LysGlyAlaAspSerLeuGluAspPheLeuTyHISgluGlyTyralaCysThrSerIle 471
QY      1711 CATGCTATTCGGACAGACAGACGGGAGCAGCTCTTGGAATTTTCCTTTGGAAG 1770
Db      472 HisGlyAspArgSerGlnArgAspArgGluGluAlaLeuHisGlnPheArgSerGlyLys 491
QY      1771 TGCCCACTTCTGTTGCTACTAGTACGTCGCCAGAGGGCTGATATGAAATGTCGA 1830
Db      492 SerProLeuLeuValAlaThrAlaValAlaAlaAlaArgGlyLeuAspIleSerAsnValLys 511
QY      1831 CATGTTATCAATTTGATCTCTCTTACCATGATGATGATGATGATGATGATGATG 1890
Db      512 HisValIleAsnPheAspLeuProSerAspIleGluGluTyValHisArgIleGlyArg 531
QY      1891 ACTGGTCTGTGGGATTAAGTACGACAGCAATTTCTTTGATCTTTGATCGATGATAC 1950
Db      532 ThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePheAsn---GluArgAsnIle 550
QY      1951 CATTTACACAGCCTCTAGTAAAGTATGACAGATGCAAGATGCTTCCGATG 2010
Db      551 AsnIleThrIysAspLeuLeuAspLeuValGluAlaLysGlnValProSerTrp 570
QY      2011 TTGGAAGAAATTCCTTTAGTACATATCTCGCTTCACTGATGATGATGATGATG 2070
Db      571 LeuGluAsnMetAlaPheGluHisHIS-----TyrLysGlySerSerArgGlyArg 587
QY      2071 GTG-----TTGCATCACTGATGATGATGATGATGATGATGATGATGATG 2106
Db      588 SerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAspTyrArgGln---SerSer 606
QY      2107 ACTTGAACACAGCTGGGTTTCTTCTTCAAGAGCT 2142
Db      607 GlyAlaSerSerSerSerPheSerSerArgAla 618

RESULT 10
ID      AAM81502 standard; Protein; 662 AA.
XX      AAM81502;
XX      AAM81502;
XX      02-FEB-1999 (first entry)
XX      XX
XX      Dead Box X (DBX) gene short transcript amino acid sequence.
XX      DE
XX      Non-recombining region; human; Y chromosome; X homologue; testis; DBX;
XX      KW
XX      Infertility; sperm; gene alteration; inhibitor; Dead Box X.
XX      OS
XX      Homo sapiens.
XX      PN
XX      MO9846747-A2.
XX      PD
XX      22-OCT-1998.
XX      PF
XX      10-APR-1998; 98MO-US07115.
XX      PR
XX      11-APR-1997; 97US-0041877.
XX      PA
XX      (WHD) WHITEHEAD INST BIOMEDICAL RES.
XX      PI
XX      Lahn BT, Page DC;
XX      DR
XX      WPI: 1998-568729/48.
XX      DR
XX      N-PSDB; AAV69632.

```

```

PT      Novel genes in the non-combining region of Y chromosome - useful to
PT      disclose if male infertility or reduced sperm count has a genetic
PT      basis
XX      XX
XX      Disclosure; Fig 3A-B; 54pp; English.
XX      XX
XX      This represents the amino acid sequence of the Dead Box X (DBX) gene
XX      short transcript. The invention relates to genes occurring on the non-
XX      recombining region of the human Y chromosome. The sequences fall into two
XX      classes: (1) X-homologous DNA which are expressed in many organs, having
XX      functional X homologues and (2) testis-specific DNA sequences. Y
XX      chromosomal DNA from males with known conditions such as infertility and
XX      reduced sperm count can be assessed using the invention to determine
XX      whether the condition is associated with or caused by the occurrence of
XX      the gene or gene alteration. Candidate inhibitors of the enzymatic
XX      activity of the genes can be assessed using in vitro assays.
XX      XX
XX      Sequence 662 AA:
XX      SQ
XX      Alignment Scores:
XX      Pred. No.: 7,48e-107 Length: 662
XX      Score: 1215.00 Matches: 293
XX      Percent Similarity: 56.64% Conservative: 91
XX      Best Local Similarity: 43.22% Mismatches: 180
XX      Query Match: 30.73% Indels: 114
XX      DB: 19 Gaps: 20
XX      US-09-714-865-15 (1-2172) x AAM81502 (1-662).
XX      QY      214 GAGTGAATTAAGCGAGATTAATACATCCACATGCTGTTGAGTGGAAAGCT--- 270
XX      Db      20 AspLeuAsnSerSerAspAsnGlnSer-----GlyGlySerThrAlaSerLysGlyArg 37
XX      QY      271 -----TTGGAACACAGAGTTTTCAAACAGCAGTTTGAAGATGGTGAAT 315
XX      Db      38 TyrIleProProHisLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56
XX      QY      316 AGCTGCTGTTTCGAGAGAGGTCTAGTAATGATCGC-----GAA 354
XX      Db      57 SerSerGlyTyrSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76
XX      QY      355 GATTAATCAACACGAGACAGAGGTTTTCAGAGAGCGGCTATGACATGAGAAATTAAT 414
XX      Db      77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer----- 90
XX      QY      415 TCAGAGCTTCAGGGCCATACAGAAAGGTGGAAGGTAGTTTC-----CGAGGTGC 468
XX      Db      91 -----GlySerArgGlyArgPheAspAspArgGlyArg 101
XX      QY      469 CGTGAAGATTTGCTAGGAAGTCCAAATATGACTTAGACCCAGACGATGTATGACG 528
XX      Db      102 SerAspTyrAspGlyIleGlySer----- 109
XX      QY      529 CGCAGCTGGGCCCTTTTGTGTTCTAGAGACGATTAAGTGACACAGTAATGCTGAT 588
XX      Db      110 -----ArgGlyAsp 112
XX      QY      589 ACTTCTCAAAGCAGAGTGCAGTGAAGT---GAACGAGTGTGTTACAAAGTTTAAT 645
XX      Db      113 -----ArgSerGlyPheGlyLysPheGluArgGlyLys-----Asn 124
XX      QY      646 GAAGAAGTATAACAGAGCTCTGGAAGAATTTCTGGAAGTCAGAAACAGAGAGAGGAA 705
XX      Db      125 SerArgTyrCysAspLysSerAspGluAspAspTyrSerLys----- 138
XX      QY      706 AGTAGTATATCTCAAGACCAAAAGTACCTACATACCCCTCTCCACCTAGAGAT--- 762
XX      Db      139 -----ProLeuProProSerGluArgLeu 146
XX      QY      763 GAGAGCTCCATCTTGGACATTTATCAGACAGGATTAATCTTGCAAAATACGACACTATT 822
XX      Db      147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166

```





XX 12-SEP-2000 (first entry)  
XX Human DBX1, an RNA helicase homologue.  
XX Topoisomerase III- $\alpha$ ; interacting protein; TIII- $\alpha$  IP; RNA helicase;  
XX DBX1; nucleosome; supercoiling; chromosome segregation; recombination;  
XX stabilisation; cell division; apoptosis; cell cycle regulation;  
XX cytotoxic; anti-tumour; DBX1.  
XX Homo sapiens.  
XX MO200032768-A1.  
XX 08-JUN-2000.  
XX 29-NOV-1999; 99MO-FR02952.  
XX 30-NOV-1998; 98FR-0015081.  
XX (AVET) AVENTIS PHARMA SA.  
XX Fournier A, Goulaouic H, Riou J;  
XX MPI; 2000-412316/35.  
XX N-PSDB: AAA29207.  
XX New nucleic acid encoding ligand for topoisomerase III  $\alpha$ , useful  
XX for inhibiting the enzyme and in drug screening, e.g. for potential  
XX anticancer agents  
XX Disclosure: Page 53-56; 68pp; French.  
XX This protein is DBX1, which shows homology with RNA helicases but the  
XX activity of a helicase has never been demonstrated and its function has  
XX not yet been identified. DBX1 possesses the 8 characteristic motifs of  
XX the helicases of the family "DEAD". In particular, it appears to be part  
XX of the sub-family represented by helicase P110. The DBX1 gene is  
XX situated on the X chromosome and its homologue, which is situated on the  
XX Y chromosome has 91 percent identity with a novel protein coding  
XX sequence. The novel protein (see AAY96482) is a human topoisomerase  
XX III- $\alpha$  interacting protein (TIII- $\alpha$  IP). TIII- $\alpha$  IP has structural  
XX features in common with RNA helicases (e.g. DBX1), which are involved in  
XX destruction of the nucleosome structure, supercoiling of DNA, segregation  
XX of newly replicated chromosomes, and recombination and stabilisation of  
XX the genome. Agents that alter interaction between TIII- $\alpha$  IP and TIII- $\alpha$   
XX therefore modulate cell division, replication, transcription,  
XX translation, splicing and DNA recombination or repair, so may slow cell  
XX growth, block the cell cycle or induce apoptosis. Antibodies and ligands  
XX of TIII- $\alpha$  IP are used to prevent, treat or alleviate diseases that  
XX involve abnormal regulation of the cell cycle, i.e. they are potential  
XX anti-tumour (cytotoxic) agents.  
SQ Sequence 662 AA;  
Alignment Scores:  
Pred. No.: 7,48e-107 Length: 662  
Score: 1215.00 Matches: 293  
Percent Similarity: 56.64% Conservative: 91  
Best Local Similarity: 43.22% Mismatches: 180  
Query Match: 30.73% Gaps: 114  
DB: 21 Gaps: 20  
US-09-714-865-15 (1-2172) x AAY96483 (1-662)  
QY 214 GAGTGTAAATGACGATATACATCCAAATGGGTGTTTGGATGGAAAGAGT--- 270  
DB 20 AspleuAenSerSerAspAsnGlnSer-----GlyGlySerThrAlaSerLysGlyArg 37  
QY 271 -----TTTGAACACGAGGTTTAAACACGAGGTTTGAAGATGGTAT 315  
DB 38 TyrIleProPheHisLeuArgSerAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56

QY 316 AGCTGTGTTTCTGAGAGAGTCTAGTAATGACTGC-----GAA 354  
DB 57 SerSerGlyTyrSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76  
QY 355 GATPATCCAAACACGAGACAGAGGTTTCCAGAGAGCGGCTTATGAGATGGAATTAAT 414  
DB 77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer----- 90  
QY 415 TCAGAGCTTCAGAGCGCATACAGAAAGAGTGAAGTGTTC-----CGAGGTGC 468  
DB 91 -----GlySerArgGlyArgPheAspAspArgGlyArg 101  
QY 469 CGTGAAGATTTGTGTAGAGAGTCCAAATATGATTAATGACCCAGACGAATGATGCAG 528  
DB 102 SerAspTyrAspGlyLysGlySer----- 109  
QY 529 CGCACTGTGGCTTTTGGTTCTAGAGACCACTATTAAAGTGCACAGTAATGGTAT 588  
DB 110 -----ArgGlyAsp 112  
QY 589 ACTTTCGAAGACAGAGTGCAGTGAAGT---GAACGAGGTGTTTCAAGTTTAAT 645  
DB 113 -----ArgSerGlyPheGlyLysPheGluArgGlyLys-----Asn 124  
QY 646 GAAGAAGTAATAACAGGCTCTGGAAGAAATTTCTTGAAGTCAAGACAGAGAGAGAGAA 705  
DB 125 SerArgTyrCysAspLysSerAspArgGluAspAspTyrSerLys----- 138  
QY 706 AGTAGTGATATCAAGAGACCAAAAGTGACCTACATACCCCTCTCCACTGAGAT--- 762  
DB 139 -----ProLeuProProSerGluArgLeu 146  
QY 763 GAGAGCTCCATCTTGGACATTTATCAGACAGGATTAACCTGCAAAATACGACACTAT 822  
DB 147 GluGlnGluLeuPheSerGlyGlyAsnThrGlyIleAsnThrGlyLysArgAspPhe 166  
QY 823 CTGTGGAAGTGTCTGACATGATGACACACGACATTTCTGACTTTGAAGAGCTAAT 882  
DB 167 ProValGluAlaThrGlyAsnAsnCysProProHisIleLeuSerPheSerAspValGlu 186  
QY 883 CTCTGTACAGACTGAATTAACAACATTGCTAAAGCTGTTATACCTAAGCTTACTCTG 942  
DB 187 MetGlyGluIleLeuMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206  
QY 943 CAATAATACAGTATTCCTATCATCTGACAGACGAGATTTGATGGCTTGCTCAACA 1002  
DB 207 GlnLysHisAlaIleProIleIleLysGluLysArgAspLeuThrAlaCysAlaGlnThr 226  
QY 1003 GGGTCTGGAAGACTGCGGCTTTCTCTTCCATCAATTTGGCTCATATGATGATGATGA 1062  
DB 227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246  
QY 1063 -----ATACTGCCAGTCTG-----TTTAAAGCTTCAGGAA 1095  
DB 247 ProGlyGluAlaLeuAlaLeuAlaMetLysGluAsnGlyArgTyrGlyArgAspLysGlnTyr 266  
QY 1096 CCAAGAGTATTTATGTAAGACCAACTCGAATGGTCAACAGATTTATTTGGAAGCC 1155  
DB 267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluGluAla 286  
QY 1156 AGAAATATTTCTTTGGAGCTGTGTGAAGAGCTGTGTTATATATGAGGGAACCCAGCTG 1215  
DB 287 ArgLysPheSerTyrArgSerArgValArgProCysValAlaTyrGlyGlyAlaAspIle 306  
QY 1216 GGACATTCAATTTGCAAAATAGTACAAAGCTGTATATATTAATGCTTACTCTGGAAGA 1275  
DB 307 GlyGlnGlnIleLeuArgAspLeuGluArgGlyCysHisIleLeuLeuValAlaThrProGlyArg 326  
QY 1276 CTGATGATATCAATAGGCAAAAGAAAGATTTGGTCTCAACAGATCAATATCTTACTGTTG 1335  
DB 327 LeuValAspMetCysGluArgGlyLysIleGlyLeuAspPheCysGlyLysIleValLeu 346  
QY 1336 GATGAAGCTGATCGCATGTTGATGATGGTTTGGTCCAGAAATGAAGAAATTATTTCT 1395

```

Db      347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgArgIleValGlu 366
QY      1396 TGCCCAAGATGTCATCAAGAAAGAAACAGCCCAACCTTATGTCATGCAACTTCCCA 1455
Db      367 GluAspThrMetProPolysGlyValArgHisThrMetCysPheSerAlaThrPhePro 386
QY      1456 GAGAAATTCAAAGTTGGCTGCAGAGTTTAAAGTCAATATATCTGTTGGTCTGTT 1515
Db      387 LysGlnIleGlnMetLeuAlaArgAspPheLeu--AspGlnTyrIlePheLeuAlaVal 405
QY      1516 GGCACAGTGGGTGAGAGATGTAGAGATTCACAGACACCGTTCCCAAGTTGGCCAGTTC 1575
Db      406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValAlaIlePvalGluGlnSer 425
QY      1576 TCAAAAAGAGAAAAGCGTTGAAATCTCGAAGAACATAGGG--GATGAAAGAACATAG 1632
Db      426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445
QY      1633 GTCCTTGTGAAACTTAAGAAAAAGCAGATTTTACTCAACTTTCTTGTCAAGAAAA 1692
Db      446 ValPheValGlnThrLysLysGlyAlaAspSerLeuLysAspPheLeuTyrHisGluGly 465
QY      1693 ATATCAACTCAAGATATCATGTGATCGGAGACAGAGACGCGGAGCAAGCTTTGGA 1752
Db      466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluAlaLeuHis 485
QY      1753 GATTTCGCTTGGCAAGTCCCGCAGTCTCTGCTGCTTCACTTCAAGTAGTCCAGAGGGCTG 1812
Db      486 GluPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505
QY      1813 GATATTTGAAATGTGCAACATGTTATCAATTTTGTCTTCTTACCATGATGAATAT 1872
Db      506 AspIleSerAsnValLysHisValIleAsnPheSerLeuProSerAspIleGlnGluTyr 525
QY      1873 GTTCATGGAATTTGGCGCTACTGCTGTTGGGAAATACGCGACGCAATTTCTTTT 1932
Db      526 ValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545
QY      1933 GATCTTGATCGGATTAACATTTTAGCACAGCTCTATTAAGATTTAGACAGATGCTCAA 1992
Db      546 Asn--GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuValGluAlaLys 564
QY      1993 CAGGATGTCCTGCATGCTTGAAGAATGCTTATGATCATCATTCCTGCTCAGT 2052
Db      565 GlnGluValProSerThrPleuGluAsnMetAlaTyrGlnHis-----TyrLys 581
QY      2053 GGTAAGTCAAGAGAAACGTG-----TTGCATCAGTTGAT 2088
Db      582 GlySerSerArgLysSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601
QY      2089 ACCGAGAAGGGCAAGACACTTTGAACACAGCTGGTTTCTTTTACAGACT 2142
Db      602 TyrArgGln--SerSerGlyAlaSerSerSerSerSerSerSerSerArgAla 618

```

RESULT 13  
AAW81503  
ID AAW81503 standard; Protein: 660 AA.

AC AAW81503;

DT 02-FEB-1999 (first entry)

DE Dead Box Y (DBY) gene product.

KM Non-recombining region; human; Y chromosome; X homologue; testis; DBY;

KW Infertility; sperm; gene alteration; inhibitor; Dead Box Y.

OS Homo sapiens.

PN W09846747-A2.

PD 22-OCT-1998.

```

XX      10-APR-1998; 98MO-US07115.
XX
XX      11-APR-1997; 97US-0041877.
PR
XX      (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX      Lahn BT, Page DC;
XX
XX      WPI: 1998-568729/48.
DR
XX      N-PSDB; AAV69633.
DR
XX
XX      Novel genes in the non-combining region of Y chromosome - useful to
PT      diagnose if male infertility or reduced sperm count has a genetic
XX      basis
XX
XX      Claim 6, Fig 3A-B, 54pp; English.
XX
XX      This represents the amino acid sequence of the Dead Box Y (DBY) gene
CC      product. The invention relates to genes occurring on the non-recombining
CC      region of the human Y chromosome. The sequences fall into two classes:
CC      (1) X-homologous DNA which are expressed in many organs, having
CC      functional X homologues and (2) testis-specific DNA sequences. Y
CC      chromosomal DNA from males with known conditions such as infertility and
CC      reduced sperm count can be assessed using the invention to determine
CC      whether the condition is associated with or caused by the occurrence of
CC      the gene or gene alteration. Candidate inhibitors of the enzymatic
XX      activity of the genes can be assessed using in vitro assays.
XX
XX      Sequence 660 AA;
XX
XX      SQ

```

Alignment Scores:

Pred. No.:	9,55E-104	Length:	660
Score:	1182.50	Matches:	278
Percent Similarity:	58.06%	Conservative:	93
Best Local Similarity:	43.51%	Mismatches:	167
Query Match:	29.91%	Indels:	101
DB:	19	Gaps:	18

US-09-714-865-15 (1-2172) x AAW81503 (1-660)

```

QY      301 TTGAAGATGATGATAGCTGCTGTTCTGAGAGAGCTTAGTAATGACTGC----- 351
Db      52 PheHisAspLysAspSerSerGlyTyrPserCysSerLysAspLysAspAlaTyrSerSer 71
QY      352 -----GAAATATATCCACACGAGACAGAGGGTTTCCAGAGAGCGGCTATCGA 402
Db      72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGlnArgLysSer----- 89
QY      403 GATGAATAATATTCAGAGAGCTTACAGAGGCTTACAGAGAGAGAGAGTAGTTTC--- 459
Db      90 -----GlySerArgGlyAlaArgPheAsp 96
QY      460 ---CGAGTTGCGTGAGAGATTGCTTAGAAGTCCAAATATATGACTTAGACCCAGAC 516
Db      97 AspArgGly-----ArgSerAspTyrAsp----- 104
QY      517 GAATGTATGACGCGCAGCTGTCCTTTTGTGTTTGAAGAACACAGTATTAAGTGGCACA 576
Db      105 -----GlyIle 106
QY      577 GGTAAATGATATCTTCTCAAGCAGAGTGGCAGTGAAGT---GAACGAGGTGGTTAC 633
Db      107 GlyAsn-----ArgGluArgProGlyPheGlyArgPheGlnArgSerGlyHis 122
QY      634 AAAGGTTTAATGAAGAAGTATATACAGGCTCTGAAAGAAATTTCTTGAAGTACAGAGCA 693
Db      123 SerArgTyrCysAspLys-----SerValGlnAspAspLys----- 136
QY      694 GAAGAGAGAGAAAGTATGATATCTCAAGACCAAAAGTACCTACATACCCCTCTCCA 753
Db      137 -----ProLeuProPro 140

```

```

QY 754 CCTGAGAT---GAGACATCCATCTTGACATTATGACAGAGCATAACTTGACAAA 810
Db 141 SerGluArgLeuGluGlnGluLeuPheSerGlyGlyLeuThrGlyIleAsnPheGluLys 160
QY 811 TAGGACACTATCTTGGAGAGTGTCTGGACATGATGACACAGCAAGATTCTGACTTTT 870
Db 161 TyrAspAspIleProValGluIleThrGlySerAsnGlyProPheIleIleGluAsnPhe 180
QY 871 GAAGAACGTAATCTCTGACAGACTGAATGAACAACATGCTTAAGCTGTATACAG 930
Db 181 SerAspIleAspMetGlyIleIleLeuGlyAsnIleGluLeuThrArgTyrThrArg 200
QY 931 CTATCTCTGTGCAAAAATACAGATATCTATCTATCTTACGAGGACGAGATTGATGAGCT 990
Db 201 ProThrProValGlnIleValHisAlaIleProIleIleGlySerGlyValAspLeuValAla 220
QY 991 TGTGCTCAACAGAGGCTCTGGAGAGACTGCGGCTTTTCTCTCAACAATTTGGCTCATATG 1050
Db 221 CysAlaGlnThrGlySerGlyLysThrAlaIlePheLeuLeuProIleLeuSerGlnIle 240
QY 1051 ATGCATGATGAATACCTGCCAGTCGTTTAAAGTTGCAGAA----- 1095
Db 241 TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgTyrGly 259
QY 1096 -----CCAGAGTGATATTATTGTACACCACTCCAGAACTGCAGAACTGCAACGAG 1140
Db 260 ArgArgLysGlnTyrProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGln 279
QY 1141 ATTATTTTGAAGCCAGAAAATTTTCTTTTGGAGCTTGTGAAGCTGTTGTTATATAT 1200
Db 280 IleTyrGlnGluAlaArgLysPheSerTyrArgSerArgValArgProCysValValTyr 299
QY 1201 GGGGAGCCCGAGCTGGGACATTCATCAATAGACAAATAGACAAAGCTGTGAATATTTATGT 1260
Db 300 GlyGlyAlaAspIleGlyGlnGlnIleArgSerLeuGluArgGlyCysHisLeuLeuVal 319
QY 1261 GCTACTCCTGGAAACAGTATGATATCATAGCAAGAAAGATTGCTCAACAGATC 1320
Db 320 AlaThrProGlyArgLeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCys 339
QY 1321 AAATACTAGTTTGGATGAAGCTGATGCAATGTTGATGCTGTTTGGTCCAGAAATG 1380
Db 340 LysTyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGlnProGlnIle 359
QY 1381 AAGAAATTAATTTCTGCCCGACGAATGCCATCAAGAAACGCCCAACCCCTATATGTC 1440
Db 360 ArgArgIleValGlnGlnAspThrMetProProGlyValArgHisThrMetMetPhe 379
QY 1441 AGTGCAACTTTCCAGAGGAATTCAAAGTTGCTGCGAGAGTTTAAAGTCAAAATAT 1500
Db 380 SerIleThrPheProLysGlnIleGlnMetLeuAlaArgAspPheLeu---AspGlnTyr 398
QY 1501 CTGTTTGTCTGTTGGACAAAGTGGGTGGACATGTAGAGATGTTTCAGCAGACCGTTCTC 1560
Db 399 IlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnIleValVal 418
QY 1561 CAATTTGCCCGAGCTCTCAAAAAGAGAAAGCTCGTGAATATTTGCCAAAATAGCG--- 1617
Db 419 TrpValGluAspLeuAspLysArgSerPheLeuLeuAspIleLeuGlyAlaThrGlySer 438
QY 1618 GATGAAGAAGACTATGAGCTTTGTTGAATCAAGAAAAGAGAGATTTTACTGCAACTTTT 1677
Db 439 AspSerLeuThrLeuValPheValGlnIleThrLysLysGlyAlaAspSerLeuGluAspPhe 458
QY 1678 CTTTGTCAAGAAAAATATCAACTACAGATTCATGCTGTGCGGAAACAGAGAGCGG 1737
Db 459 LeuTyrHisGlnGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478
QY 1738 GAGCAAGCTCTTGAGAGATTTTCGCTTGGAGAAAGTCCAGATTTCTTGCTACTTCAGTA 1797
Db 479 GlnGluAlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaVal 498
QY 1798 GCTGCCAGAGGGCGATATTTGAATAATGTGCAACATGTTATCAATTTTGATCTTCCTTCT 1857

```

```

Db 499 AlaAlaArgGlyLeuAspIleSerAsnValArgHisValIleAsnPheAspLeuProSer 518
QY 1858 ACCATTGATGAATATGTTCTCATTCGAATTTGGCGCTACTGCTGTGGATATCTGCAGCA 1917
Db 519 AspIleGlnGluTyrValHisArgIleGlyArgThrGlyArgValGlnAsnLeuGlyLeu 538
QY 1918 GCATTTCTCTTTTGTGATCTGATCGATACCATTTAGCACAGCGCTCTAGTAAAGTA 1977
Db 539 AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557
QY 1978 TTGACAGATGCTCAACAGATGTTCTGATGCTGTGGAAGAAATTCGCTTTATACATAC 2037
Db 558 LeuValGlnLysGlnGlnIleValProSerTrpLeuGlnAsnMetAlaTyrGlnHisHis 577
QY 2038 ATTCCCTGCTTCACT---GGTACTACAGAGAAAGCTGTTTCA-----TCA 2082
Db 578 TyrLysGlyLysSerArgLysSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597
QY 2083 GTTGATACAGAAAGGCAAGACACTTTGACACAGCTGCTGTTTCTTCTTCACGA 2139
Db 598 ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615

RESULT 14
AAG17433
ID AAG17433 standard; Protein; 612 AA.
AC AAG17433;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18449.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EPI033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.

```

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140685.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142350.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 21-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 25-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.:	8,79e-96	Length:	612
Score:	1099.00	Matches:	252
Percent Similarity:	57.97%	Conservative:	97
Best Local Similarity:	41.86%	Mismatches:	167
Query Match:	27.79%	Indels:	86
DB:	21	Gaps:	15

US-09-714-865-15 (1-2172) x AAG17433 (1-612)

QY 313 GATAGCTGCTGTTCTGAGAGAGCTGATGATGACGCGAATGATTCGAAACAGCGGAC 372  
 ||| :|||: :||| :|||:  
 Db 43 Asph1salaglytyrglyglylnProalaglysertrgrtpralaproproserSergly 62  
 QY 373 AGAGGTTTTCCAAAGAGCGGCTATCGAGATGAAATTAATTCAGAACTTCAGGCGCA 432  
 ||| :|||: :||| :|||:  
 Db 63 GlyglylalaSerglyglytyrArgSanaSp----- 74  
 QY 433 TACGAAGAGGTGGAAGGTGATGTTCCAGGTTCCGCGAGGATTTGGTCTAGSAGT 492  
 ||| :|||: :||| :|||:  
 Db 75 -----GlyglylArgThrGlytyr-----GlytyrGlylalaGly--- 85  
 QY 493 CCAATATGACTTAGACCCAGACGAATGTATCGACGCGCACTGGTGGCTTTTGGTTCT 552  
 ||| :|||: :||| :|||:  
 Db 86 -----Glyglyl-----Glyglyl----- 87  
 QY 553 AGAAGACCAGTATTAACTGCGACAGGTAACTGTAATCTTCTCAAGACAGGTGGAGT 612  
 ||| :|||: :||| :|||:  
 Db 88 -----GlyglylGlyglyl-----GlyTyr 94  
 QY 613 GGAAGTGAACGAGTGTGTTCAAAAGTTTAAATGAAGATTAATACAGCTCTCGAAAG 672  
 :||| :|||: :||| :|||:  
 Db 95 AsnAsnArgSerglytyrPaspArgArgGluArgGluVal----- 108  
 QY 673 AATCTTGAAGTGAAGAGAGAGAGAGAGAAAGTAGTAGTACTCAAGACCAAAAGTG 732  
 ||| :|||: :||| :|||:  
 Db 109 AsnProPheglylaSpAspAlaGlu----- 116  
 QY 733 ACCTACATACCCCTCTCTCCACCTGAGAGTAGAGACTTCATCTTTCACATTATACACA 792  
 :||| :|||: :||| :|||:  
 Db 117 -----LeuGluProValPheThrGluGlnGluAsn-----Thr 127  
 QY 793 GGCATTAACCTTCGCAAAATAGACACTTCTCTGGAAGTGTGGAATGACATGACACA 852  
 ||| :|||: :||| :|||:  
 Db 128 GlylleAsnPheAspAlaTyrGluAspIleProValGluThrSerglyGlyAspValPro 147  
 QY 853 CCACCAATTTCTGACTTTGGAAGAGCTATCTCTCTGACAGACTGTAATACAGATGCT 912  
 ||| :|||: :||| :|||:  
 Db 148 ProProValaAsnThrPheAlaAspIleAspIleuGlyAspAlaAlaAsnLeuAsnIleArg 167  
 QY 913 AAAGCTGTATTAATTAAGCTTACTCTCTGCAAAAATACAGATTTCTTATATCTTGA 972  
 :||| :|||: :||| :|||:  
 Db 168 ArgCysLysTyrValArgProThrProValGlnArgHisAlaIleProIleLeuLeuAla 187  
 QY 973 GGAGCAGATTTGAGCGCTTGTGCTCAACAGAGCTGGGAAGACTGGCGCTTCTCTCTA 1032  
 ||| :|||: :||| :|||:  
 Db 188 GluArgAspLeuMetAlaCysAlaGlnThrGlySerglyLysThrAlaAlaPheCysPhe 207  
 QY 1033 CCAATTTGGCTCATGATGATCATGAT---GGAATTAACCTGCGAGCTTTTAAAGAGTTG 1089  
 ||| :|||: :||| :|||:  
 Db 208 ProIleIleSerglylleMetLysAspGlnHisValGluArgProArgGlySerArgAla 227  
 QY 1090 CAGAACCAAGAGTGTATTATTGTAGACCAACTCGAGATTTGGTCAACAGATTTATTG 1149  
 ||| :|||: :||| :|||:  
 Db 228 ValTyrProPheAlaValIleLeuSerProThrArgGluLeuAlaCysGlnIleHisAsp 247

QY 1150 GAAGCCAGAAAAATTTCTTTGGGACTGTGTGAAGACTGTGTATATATATGGGGAGCC 1209  
 ||| :|||: :||| :|||:  
 Db 248 GluAlaLysLysPheSerytyrGlnThrGlyValHisValValAlaTyrGlytyr 267  
 QY 1210 CAGCTGGACATTCATTAATGACAAATAGTACAAAGCTGTGAATATATTATGTGCTACTCT 1269  
 :||| :|||: :||| :|||:  
 Db 268 ProIleHisGlnGlnLeuArgGluArgGlyCysAspIleLeuValAlaThrPro 287  
 QY 1270 GGAAGCTGATGATGATTCATAGCCAGAAAAAGATGTGCTCAACAGATTCATTAATCTTA 1329  
 ||| :|||: :||| :|||:  
 Db 288 GlyArgLeuAsnAspLeuGluArgAlaArgValSerMetGlnMetIleArgPheLeu 307  
 QY 1330 GTTTGGATGAAGCTGATGCGATGTTGATATGCTTTTGGTCCAGAAAAAGAAAGTTA 1389  
 ||| :|||: :||| :|||:  
 Db 308 AlaLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgLysIle 327  
 QY 1390 ATTTCTTGGCCAGGAATGCCATCAAGAAACAGCCGCAACCTTATGTTCACTGCAACT 1449  
 :||| :|||: :||| :|||:  
 Db 328 ValGluGlnMetAspMetProProArgGlyValArgGlnThrMetLeuPheSerAlaThr 347  
 QY 1450 TTTCCAGAGGAATTCAAAGGTTGGCTGCGAGAGTTTAAAGCAATTAATCTGTTGTT 1509  
 ||| :|||: :||| :|||:  
 Db 348 PheProSerylIleGlnArgLeuAlaAlaAspPheMet---SerAsnTyrIlePheLeu 366  
 QY 1510 GCTGTTGACAACTGGGTGGAGCATGTAGATGTTCAAGACACGTTCTCAAGTTGGC 1569  
 ||| :|||: :||| :|||:  
 Db 367 AlaValGlyArgValGlySeryThrAspLeuIleThrGlnArgValGluPheValGln 386  
 QY 1570 CAGTTCTCAAAAAGAAAAAGCTGTTGAAATTCG-----CGAAACATAGGGGAT 1620  
 :||| :|||: :||| :|||:  
 Db 387 GluSerAspLysArgSeryHisLeuMetAspLeuLeuHisAlaGlnArgGluThrGlnAsp 406  
 QY 1621 GAAGA-----ACTATGCTTGTGTTGAACATAGAAAAAGACAGATTTTACTGCAACT 1674  
 :||| :|||: :||| :|||:  
 Db 407 LysGlnSeryLeuThrLeuValPheValGluThrLysArgGlyAlaAspThrLeuGluAsn 426  
 QY 1675 TTTCTTGTTCAGAAAAATATATCACTACAAGTATCCATGATGCGGAACAGAGAG 1734  
 ||| :|||: :||| :|||:  
 Db 427 TyrLeuCysMetAsnGlnPheProAlaThrSeryIleHisGlyAspArgThrGlnGln 446  
 QY 1735 CGGAGACAGCTCTTGGAGATTTTCGCTTGGAAAGTCCCACTTCTTGTGCTACTTCA 1794  
 :||| :|||: :||| :|||:  
 Db 447 ArgGluValAlaLeuArgSeryPheLysThrGlyArgThrProIleLeuValAlaThrAsp 466  
 QY 1795 GTAGCTGCCAGAGGCTGGAATTTGAATGCAACATGTTTCAATTTGATCTTCTCT 1854  
 ||| :|||: :||| :|||:  
 Db 467 ValAlaAlaArgGlyLeuAspIleProHisValAlaHisValAlaAsnPheAspLeuPro 486  
 QY 1855 TCTACCATGATGAATATGTTTCATCGAATTTGGCGTACTGGTCTTGGAAATACTGGC 1914  
 :||| :|||: :||| :|||:  
 Db 487 AsnAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyLysSergly 506  
 QY 1915 AGAGCAATTTCTTTTGTGATCTTGAATCGGATTAACCATTTAGACAGCTCTTAGTAAA 1974  
 ||| :|||: :||| :|||:  
 Db 507 IleAlaThrLalaPhePheAsn---GluAsnAsnAlaGlnLeuAlaAspSeryLeuAlaGlu 525  
 QY 1975 GTATTCGACATGCTCAACAGAGATTCCTGCATGTTGGAAGAATATGCTTTAGTACA 2034  
 :||| :|||: :||| :|||:  
 Db 526 LeuMetGlnGluAlaAsnGlnGluValProGluThrPheuThrArgTyrAlaSerArgAla 545  
 QY 2035 TACATTCCTGCTTCACTAGTGTAGTACAGA---GGAACGCTGTTTGCATCAGTTGATACC 2091  
 ||| :|||: :||| :|||:  
 Db 546 SerPheGlyGly-----GlyLysLysArgSeryGlyLysArgPheGlyArgAspPhe 563  
 QY 2092 AGAAAG 2097  
 ||| :|||: :||| :|||:  
 Db 564 ArgArg 565  
 RESULT 15  
 AAG20655  
 ID AAG20655 standard; Protein: 612 AA.  
 XX  
 AC AAG20655;



XX 17-OCT-2000 (first entry)  
DT Arabidopsis thaliana protein fragment SEQ ID NO: 22931.  
XX  
DE  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridization assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX EPI033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143472.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148584.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161921.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Alignment Scores:  
 Pred. No.: 8.79e-96 Length: 612  
 Score: 1099.00 Matches: 252  
 Percent Similarity: 57.97% Conservative: 97  
 Best Local Similarity: 41.86% Mismatches: 167  
 Query Match: 27.79% Indels: 86  
 DB: 21 Gaps: 15

US-09-714-865-15 (1-2172) x AAG20655 (1-612)

QY 313 GATAGCTTGTTGGTGGAGAGCTTATGACTCGGAAGATATCCAGACGGAAC 372  
 DB 43 Asph1salalaglyTYgLYgLYglnProAlaeglyserArgrTalaProPioSerSerSergly 62  
 QY 373 AGAGGCTTCCAGAGAGGAGGCGATCGAGATGGAATTAATTCAGAGCTTCAGGGCCA 432  
 DB 63 GlylglygLYalaserGlylglygLYTYrArghAnsp----- 74  
 QY 433 TACAGAGAGGTGGAAGAGTAGTTCCGAGGTGCCGTGAGAGATTGTCCTAGAGAGT 492  
 DB 75 -----GlygLYArgrThrgLYTYr-----GlyTYrGlyAlaGly--- 85

QY 493 CCAATATATGACTTAGACCCAGACGAATGTATGACGCGACTGGCCCTTTTGGTCT 552  
 DB 86 -----GlygLY----- 87  
 QY 553 AGAAGACAGATTAAGTGACAGGTAAAGTGTACTTCCAAAGACGAAGTGCAGT 612  
 DB 88 -----GlygLYgLYgLYgLY-----GlyTYr 94  
 QY 613 GGAAGTGAACGAGTGGCTTCAAAAGTTAAATGAAGATTAATACAGGCTCTGGAAG 672  
 DB 95 AsnAsnArgrSerGlygLYTYrPaPaArgrGlyUAArgGlyVal----- 108  
 QY 673 AATTCCTGAGAGTCAGACAGAGAGAGGAAGAAAGTAGATATCAAGACCAAAAGTG 732  
 DB 109 AsnProPheGlyAspAspAlaGlu----- 116  
 QY 733 ACCTACATACCCCTCTCCCTCAGATGAGAGACTCCATCTTGCACATATACAGCA 792  
 DB 117 -----LeuGlnProValPheThrgInGlnGluAsn-----Thr 127  
 QY 793 GGCATAACTTGCACAAATACGACACTATCTTGGAGAGTGTCTGGACATATGACACA 852  
 DB 128 GlylLeAsnPhesAspAlaTYrGlnAspIleProValGlnThrSerGlygLYAspValPro 147  
 QY 853 CCAGCAATCTGACTCTTGGAGAGCTTAATCTCTGCAGACACTGAATAACAATTTGCT 912  
 DB 148 ProProValAsnThrPheAlaAspIleAspLeuGlyAspAlaLeuAsnLeuAsnIleAsg 167  
 QY 913 AAGCTGTTATCTAAGCTTACTCTGTGCAGAAAATACAGATTTCTATCTACTTGA 972  
 DB 168 ArgCylLeTYrValArgrProThrProValGlnArghAlaIleProIleLeuValAla 187  
 QY 973 GAGAGATTTGATGGCTGTGCTCAACAGGCTGGGAAGCTCGGGCTTTCCTCA 1032  
 DB 188 GluArAspLeuMetAlaCysAlaGlnThrgLYSerGlyLYrThrAlaAlaPheCysPne 207  
 QY 1033 CCAATTTGGCTCATATGATGATGATGAT---GGAATACCTCCAGCTGTTTAAAGAGTTG 1089  
 DB 208 ProIleIleSerGlylLeuMetLYsAspGlnHisValGlnArgrProArgrGlySerArGAla 227  
 QY 1090 CAGAACCCAGAGTGTATTTATGTAGACCAACTCGAAGATTTGTCAACGATTTATTG 1149  
 DB 228 ValTYrProPheAlaValIleLeuSerProThrArgrGlnLeuAlaCysGlnIleHisAsp 247  
 QY 1150 GAAGCCAGAAAATTTCTTTTGGAGCTTGTGTAAAGCTTGTATATATGGGGGAAC 1209  
 DB 248 GluAlaLYsLYsPheSerTYrGlnThrgLYValLYsValValAlaTYrGlygLYThr 267  
 QY 1210 CAGCTGGACATTCATTCAGCAAAATGTACAGAGCTGTAAATATATATGTCTACTCCT 1269  
 DB 268 ProIleHisGlnGlnLeuAlaArgGlnLeuGlnArgrGlyCysAspIleLeuValAlaThrPro 287  
 QY 1270 GGAAGACTGTATATCATATGAGCAAGAAAGAAAGATTGGTCTCAACAGATCAATACTTA 1329  
 DB 288 GlyArgrLeuAsnProPheLeuGlnArGAlaArgrValSerMetGlnMetIleArgrPheLeu 307  
 QY 1330 GTTTTGATTAAGCTATGCCATGTGGATATATGAGTTTGTCTCAGAAATGAAGAGTTA 1389  
 DB 308 AlaLeuAspIleAlaAspArgMetLeuAspMetGlyPheGlnProGlnIleArgrLYsIle 327  
 QY 1390 ATTTCCTGCCAGAAATGCCATCAAGAGACAGAGCCAAACCTTATGTCAGTCAACT 1449  
 DB 328 ValGlnGlnMetAspPheProProArgrGlyValArgrGlnThrMetLeuPheSerAlaThr 347  
 QY 1450 TTTCCAGAGAAATTCAAAGTTGGCTGACAGCTTTTAAAGTCAATATCTGTTGTT 1509  
 DB 348 PheProSerGlnIleGlnArgrLeuAlaAlaAspPheMet---SerAnTYrIlePheLeu 366  
 QY 1510 GCTGTTGACAGAGTGGGTGAGCATGTAGAGATGTTTCAGCAGACCGTTCTCCAAGTTGCC 1569  
 DB 367 AlaValGlyArgrValGlySerSerThrAspLeuIleThrGlnArgrValGluPheValGln 386



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 06:09:16 ; Search time 2771 Seconds  
(without alignments)  
12694.550 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 2172  
Sequence: 1 acgggagatgaatgcattggga.....tagatgatgcattgcattggat 2172

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2150.4	99.0	2523	11	BC030638 Homo sapi
2	1528.4	70.4	2759	11	AK014844 Mus muscu
3	683	31.4	798	12	BG717110 602689191
4	640.8	29.5	822	13	BI464267 603203379
5	631.8	29.1	792	14	BQ429126 AGENCOURT
6	624.8	28.8	747	12	BG719449 602690172

7	620.4	28.6	682	12	BG718234	BG718234	602696209
8	620.4	28.6	693	13	BI560849	BI560849	603254018
9	615.6	28.3	830	13	BI462207	BI462207	603205520
10	612.4	28.2	918	13	BI561001	BI561001	603254391
11	610.4	28.1	846	12	BG718634	BG718634	602686739
12	597.4	27.5	699	12	BG717822	BG717822	602720504
13	596.6	27.5	758	13	BI559479	BI559479	603252992
14	594.8	27.4	861	13	BI461393	BI461393	603206861
15	593.4	27.3	657	12	BG717812	BG717812	602693829
16	584.6	26.9	743	13	BI463502	BI463502	603203596
17	581.8	26.8	884	13	BI560499	BI560499	603254476
18	564.4	26.0	651	12	BG717354	BG717354	602689688
19	559.8	25.8	700	12	BG718861	BG718861	602699337
20	544.4	25.1	847	13	BI562624	BI562624	603256559
21	542.6	25.0	810	13	BI463846	BI463846	603202572
22	524.8	24.2	841	13	BI459115	BI459115	603199485
23	504	23.2	917	13	BI463447	BI463447	603204483
24	483.6	22.3	546	13	BM539751	BM539751	603204483
25	449.8	20.7	670	14	BQ443738	BQ443738	UI-M-EMO-
26	433.8	20.0	482	9	AL042306	AL042306	DFEP2434M
27	420.2	19.3	780	10	AM153705	AM153705	fi24605.Y
28	389	17.9	726	12	BF611759	BF611759	de88402.Y
29	371	17.1	4626	11	BC007668	BC007668	Homo sapi
30	368.8	17.0	604	13	BI460628	BI460628	603201271
31	339	15.6	632	9	AL785068	AL785068	AL785068
32	333	15.3	335	9	AA383535	AA383535	EST96928
33	310.4	14.3	555	9	AA399611	AA399611	FE93407.S
34	298.4	13.7	491	14	BQ520442	BQ520442	NISC.n104
35	296.2	13.6	537	14	BI083771	BI083771	UI-M-ANI-
36	290.8	13.4	430	9	AI184371	AI184371	qf47d11.X
37	288.4	13.3	555	9	AI217144	AI217144	qf47d11.X
38	270	12.4	533	10	AM634656	AM634656	BI23402.W
39	258.4	11.9	1084	13	BM462228	BM462228	AGENCOURT
40	247.6	11.4	976	17	CNS071VS	CNS071VS	AL425448
41	237.4	10.9	457	13	BU073747	BU073747	clone BAO
42	236.6	10.9	819	9	AA791735	AA791735	v008008.X
43	235.4	10.8	347	9	AI953070	AI953070	wq49h07.X
44	222.6	10.2	637	13	BI461238	BI461238	603206562
45	220.2	10.1	816	9	AU132682	AU132682	AU132682

## ALIGNMENTS

RESULT 1  
LOCUS BC030638 2523 bp mRNA linear HTC 21-MAY-2002  
DEFINITION Homo sapiens, VASA protein, clone IMAGE:4822528, mRNA.  
ACCESSION BC030638  
VERSION BC030638.1 GI:21040391  
KEYWORDS HTC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2523)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (20-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabbs-rc@mail.nih.gov](mailto:cgabbs-rc@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshikuni and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 33 Row: d Column: 8  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA 91: 9507236  
 This clone has the following problem: frame shifted.

## FEATURES

source

location/Qualifiers

1..2523  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /db\_xref="LocusID:54514"  
 /clone="IMAGE:4822528"  
 /issue\_type="Testis"  
 /clone\_id="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"

BASE COUNT 780 a 437 c 616 g 690 t  
 ORIGIN

Query Match 99.0%; Score 2150.4; DB 11; Length 2523;  
 Best Local Similarity 99.7%; Pred. No. 0;

Matches 2165; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

1 ATGGGGGATGAAGATTGGGAAGCAGAAATCAACCCCTCATATGCTTCTTAATGCCATA 60  
 98 ATGGGAATGAAGATTGGGAAGCAGAAATCAACCCCTCATATGCTTCTTAATGCCATA 157  
 61 TTGAGAAGATAGATTCTTGAGAAATGAGACAAATTTTAAACAGAGCTCCAGCTTCA 120  
 158 TTGAGAAGATAGATTCTTGAGAAATGAGACAAATTTTAAACAGAGCTCCAGCTTCA 217  
 121 TCATCAAAATAGATGATGACCTTCTCGAAGAGATATTTCAAGAAAGTGAATTTGCC 180  
 218 TCATCAAAATAGATGATGACCTTCTCGAAGAGATATTTCAAGAAAGTGAATTTGCC 277  
 181 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGAGTGAATTAAGCAGATTAATCATCC 240  
 278 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGAGTGAATTAAGCAGATTAATCATCC 337  
 241 ACAATGGGTGTTTGGAGTTGAAAAGATTGAAAACAGAGTTTTCAAACAGCAGG 300  
 338 ACAATGGGTGTTTGGAGTTGAAAAGATTGAAAACAGAGTTTTCAAACAGCAGG 397  
 301 TTGGAATGCTGATCTGCTGCTTCTGAGAGAGCTGATATGATCGAGAGATTAAT 360  
 398 TTGGAATGCTGATCTGCTGCTTCTGAGAGAGCTGATATGATCGAGAGATTAAT 457  
 361 CCAACACGGAACAGAGGTTTTCAGAGAGGCGGCTATCGAGATGGAATTAATTCAGAA 420  
 458 CCAACACGGAACAGAGGTTTTCAGAGAGGCGGCTATCGAGATGGAATTAATTCAGAA 517  
 421 GCTTCAGGGCCATACAGAAAGTGAAGAGTATGTTCCGAGGTTCCCGTGAAGATT 480  
 518 GCTTCAGGGCCATACAGAAAGTGAAGAGTATGTTCCGAGGTTCCCGTGAAGATT 577  
 481 GGTCTAGGAAGTCCAAATATGACTTAGACCCAGACCAATGTATGACGCCGACTGGGC 540  
 578 GGTCTAGGAAGTCCAAATATGACTTAGACCCAGACCAATGTATGACGCCGACTGGGC 637  
 541 CTTTTCGTTCTAGAGACCAAGTATTAAGTGGCAGAGTAAATGATGATCTTCAAGC 600  
 638 CTTTTCGTTCTAGAGACCAAGTATTAAGTGGCAGAGTAAATGATGATCTTCAAGC 697  
 601 AGAAGTGGCAGTGAAGTGAACGAGTGTGTTCAAAAGTATTAATGAAGAGTAAATACA 660  
 698 AGAAGTGGCAGTGAAGTGAACGAGTGTGTTCAAAAGTATTAATGAAGAGTAAATACA 757  
 661 GGTCTAGGAAGTCTTGAAGTCAAGAGCAGAGAGAGAGAAATGATGATCTCA 720  
 758 GGTCTAGGAAGTCTTGAAGTCAAGAGCAGAGAGAGAGAAATGATGATCTCA 817

721 GGACCAAAAGTACCTACATACCCCTCTCCACCTGAGAGTGAAGCTTCATCTTTGCA 780  
 818 GGACCAAAAGTACCTACATACCCCTCTCCACCTGAGAGTGAAGCTTCATCTTTGCA 877  
 781 CATTATCAGACGAGCATTAACCTTCCAGAAATACGACTATTTCTTGGAGTGTCTGGA 840  
 878 CATTATCAGACGAGCATTAACCTTCCAGAAATACGACTATTTCTTGGAGTGTCTGGA 937  
 841 CATTATCAGACGAGCATTAACCTTCCAGAAATACGACTATTTCTTGGAGTGTCTGGA 900  
 938 CATTATCAGACGAGCATTAACCTTCCAGAAATACGACTATTTCTTGGAGTGTCTGGA 997  
 901 AACCAATGCTTAAGCTGTTATATCACTTACTCTGTGCAAAAATACAGATTTCT 960  
 998 AACCAATGCTTAAGCTGTTATATCACTTACTCTGTGCAAAAATACAGATTTCT 1057  
 961 ATATATCTTGCAGAGCAGAGATTGATGCTTGTCTCAAAACAGGCTTGGAAAGCTGG 1020  
 1058 ATCATATCTTGCAGAGCAGAGATTGATGCTTGTCTCAAAACAGGCTTGGAAAGCTGG 1117  
 1021 GCTTTCCTCCTACCAATTTTGGCTCATATGATGATGAGATTAATCTGCAAGTCTTTT 1080  
 1118 GCTTTCCTCCTACCAATTTTGGCTCATATGATGATGAGATTAATCTGCAAGTCTTTT 1177  
 1081 AAGAGTTGCAGAGCAGAGTGTATTTATGAGACCAACTCGAGAAATGTCACACAG 1140  
 1178 AAGAGTTGCAGAGCAGAGTGTATTTATGAGACCAACTCGAGAAATGTCACACAG 1237  
 1141 ATTTATTTGGAAGCCGAAATTTTCTTTGGGACTTGTGTAAAGCTGTGTATATAT 1200  
 1238 ATTTATTTGGAAGCCGAAATTTTCTTTGGGACTTGTGTAAAGCTGTGTATATAT 1297  
 1201 GGGGGAACCCAGCTGGGACATTCATTCGACAAATGTCACAGGCTGTATATATATAT 1260  
 1298 GGGGGAACCCAGCTGGGACATTCATTCGACAAATGTCACAGGCTGTATATATATAT 1357  
 1261 GCTACTCTTGAAGACTGATGATATCATAGGCAAAAGAAAGATTTGCTCAAAACAGATC 1320  
 1358 GCTACTCTTGAAGACTGATGATATCATAGGCAAAAGAAAGATTTGCTCAAAACAGATC 1416  
 1321 AATATCTTATTTTGGATGAGTGTATGCTGATGTTGGATGTTTGGTCCAGAAATG 1380  
 1417 AATATCTTATTTTGGATGAGTGTATGCTGATGTTGGATGTTTGGTCCAGAAATG 1476  
 1381 AAGAGTTAATTTCTTCCAGAGAAATTCAAAGTTGGCTGACAGATTTTAAAGTCAATAT 1440  
 1477 AAGAGTTAATTTCTTCCAGAGAAATTCAAAGTTGGCTGACAGATTTTAAAGTCAATAT 1536  
 1441 AGTCAACTTTTCCAGAGAAATTCAAAGTTGGCTGACAGATTTTAAAGTCAATAT 1500  
 1537 AGTCAACTTTTCCAGAGAAATTCAAAGTTGGCTGACAGATTTTAAAGTCAATAT 1596  
 1501 CTGTTTGTGCTGTTGCAAGAGTGGTGAAGCATGTAGAGATTTTGAAGACCGTTCTC 1560  
 1597 CTGTTTGTGCTGTTGCAAGAGTGGTGAAGCATGTAGAGATTTTGAAGACCGTTCTC 1656  
 1561 CAAAGTGGCCAGTTCCAAAAGAGAAAGCTGTTGAATTTCTGCAAAATATAGGGGAT 1620  
 1657 CAAAGTGGCCAGTTCCAAAAGAGAAAGCTGTTGAATTTCTGCAAAATATAGGGGAT 1716  
 1621 GAAAGAACTATGCTTTTGTGTAATTAAGAAAGAAAGAGAGATTTTAACTCAATTTTCTT 1680  
 1717 GAAAGAACTATGCTTTTGTGTAATTAAGAAAGAAAGAGAGATTTTAACTCAATTTTCTT 1776  
 1681 TGTCAAGAAAGAAATATCACTACAGTATTCATGTGATCGGGAACAGAGAGCGGAG 1740  
 1777 TGTCAAGAAAGAAATATCACTACAGTATTCATGTGATCGGGAACAGAGAGCGGAG 1836  
 1741 CAAGCTCTTGAAGATTTTGGCTTGGAAAGTCCAGTCTTGTGCTACTTCAATAGT 1800  
 1837 CAAGCTCTTGAAGATTTTGGCTTGGAAAGTCCAGTCTTGTGCTACTTCAATAGT 1896  
 1801 GCAGAGGCTGGAATTTGAAATATGCAACATGTTATCAATTTTGTATCTTCTTAC 1860

[illegible]

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE
Aizawa, T., Hara, A., Fukunishi, Y., Komoto, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondou, S., Yamakata, I., Salto, T., Okazaki, Y., Gojobori, T., Bono, H., Kasubuchi, T., Saito, R., Kadote, K., Matsuda, H., Ashburner, M., Batalov, S., Casavet, T., Fleischmann, W., Gaasterland, P., Gissi, C., King, B., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Bottelli, D., Bojunga, N., Caminici, P., de Bonaldo, M.F., Brownstein, M.J., Balt, C., Fletcher, C., Fujita, M., Gariboldi, M., Gusticich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, T., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokawa, K., Wang, K.H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)	21085660 11217851	5 (bases 1 to 2759)	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Butt, C., Caminici, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F., Hume, D., Imocani, K., Ishii, Y., Itoh, M., Itawa, M., Kaikawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koysa, S., Kuithera, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Salto, H., Salto, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Schriml, L., Shibaata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yamakata, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.  Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details.  cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in RIKEN contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGACGACGATCCAGACCTTTTTCCTTTTTTTTN 3'], cDNA was prepared by using trenaiose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGACGATTCGAGTTAATTAAATATGCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamI. Host: DH10B.  location/Qualifiers 1..2759 organism="Mus musculus" strain="C57BL/6J" db_xref="FANTOM_DB:4921508G22" db_xref="MCD:MGI:1910815" db_xref="taxon:10090" clone="4921508G22" sex="male" tissue_type="testis" clone_id="RIKEN full-length enriched mouse cDNA library" dev stage="adult"

```
gene
1. .2759
/ gene="Ddx4"
/ db_xref="GI:12852922"
/ gene="Ddx4"
/ note="DEAD (aspartate-glutamate-alanine-aspartate) box
polyptide 4
data source:MGB, source key:MGI:102670, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB29578.1"
/db_xref="GI:12852922"
/db_xref="MGI:102670"
/translation="MGDEWBAELIKPHVSSVYVPEFKDYSSGANGDTFNRTSASSD
IGSSKKEKSTITGFGKGFNGRFLNKKFEKGGSSKESNNCEONRTSRF
KRGCGKODNDSEASGPRRGSGFRGCGRGLGRPNSESDDOGTORGGLFGSR
KPAASDSGNDTYSRSGSGRGKYLNEVYTSKSKSETEGSESSDSQPKYT
YIPPEDEDSITFAHYQTGINFDKDTILVEVSGHDAFPAILTFEEANLCQTLNNT
AKAGYKTLIPVQKSLIPILVAGDLMAQOVSQKTAFLPLIAHMMRDGIVASRK
ELQPECTIVAPRELINQIYLEARKFSFGCVAVIYGGTQFSCVROICNLI
CATGRMLMDIIGKEKIGIKQYKVLVLEADMLDMGKPEMKLISCPMSEKOROT
LIPSATPPEIQRADGPKLSYLFVAIVGVGACRPOCTIILQVGYSKREKVELT
RNIGDERTWPFVETKKKADFTATFLCOEKSTISIHGDREREOALGDPKCKCV
LVATSVARGLDIENVQVINFPLSTIDEVHRIGRGCGNTGRAISFFDDSDNH
LAQPLVIVLSDAQODVPAMLEIFSTYVPSFSSSTRGAVFASVDRKNYQKHTL
NTAGISSQAPNPVDESDSWD"
polya_signal
2734. .2739
/ gene="Ddx4"
/ note="putative"
2759
/ gene="Ddx4"
/ note="putative"
polya_site
BASE COUNT 836 a 483 c 653 g 787 t
ORIGIN
Query Match 70.4%; Score 1528.4; DB 11; Length 2759;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 256; Indels 21; Gaps 4;
QY 168 AAGTGAATTTGCCCTGGCGGGAATTTTGAAACAGAGATCGTGATGTAATAACG 227
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 134 AAATGAGACACTTTTAAACAGACTTCAGCTTCATCAGATATTGCGAGTCTAGTAA 193
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 228 AGATAATACATCCAAATGGGTGTTTGAATGGAAGAAGTTTGGAAAACAGAGTTT 287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 AGAACAACATCTCAACTGCTGCTTTGGAAGAAGAAAGGCTTTGGAACAAGTTT 253
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 288 TTCAAACAGCAGTTTGAAGATGCTGATGCTCTGCTTTCTGAGAGAGTCTAGTA 347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 254 TTTAAATTAACAAGTTTGAAGAAGGATGATGCTGCTGTTCTGGAAGAAGTCT 313
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 348 CTGGGAAGATTAATCAACAGGAACAGAGGTTTCCAAAGAGAGCGGCTATCGAG 407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 314 CTGGAAGATTAATCAACTCAAGAGAGGTTTCCAAAGAGAGTGGCTCCAAAGTGG 373
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 408 AAATAATTCAGAAAGCTTCAGGCAATACAGAAAGAGTGAAGAGTATGTTCCG 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 374 AAAGATTCAAGAGATCAGGCCCTTCAGAAAGAGGGAAGAGCAAGTTCCGAG 433
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 468 CCGTGAAGAGTTTGGTCTAGAGAGTCAATAATGACTTAGACCCAGACGAATG 527
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 434 CCGTGAAGAGTTTGGTCTAGAGAGTCAATAATGACTTAGACCCAGACGAATG 493
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 528 GCGCACTGGTGGCTTTTGGTCTAGAGAGTCAATAATGACTTAGACCCAGAGT 587
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 494 GCGTGGTGGTGGCTTTTGGTCTAGAGAGTCAATAATGACTTAGACCCAGAGT 553
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 588 TACTTCTCAAGAGAGAGTGGCAGTGAAGTGAACGAGTGTGTTACAAAGTTTAA 647
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 554 CACTTACCAAGAGCA-----AGTGAAGTGGTGAAGTGTGTTACAAAGTTTAA 607
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 648 AGAAGTAAATTAACAGGCTTGAAGAGATTTCTTGAAGTCAAGAGCAAGAGGAAG 707
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 608 AGAAGTAAATTAACAGGCTTGAAGAGATTTCTTGAAGTCAAGAGCAAGAGGAAG 667
```

```
QY 708 TAGTGATTAACAAGACAAAAGTGAACCTACATACCCCTCTCCACTGAGATGAGA 767
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 668 CAGTGATAGCAAGGTCCAAAAGTGAATATATACCCCTCTCCACAGAGATGAGA 727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 768 CTCATCTTTTGAACATTAATGACAGGCAATAACTTGACAAATATGACATATCT 827
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 728 CTCATCTTTTGAACATTAATGACAGGCAATAACTTTGATTAATATGACATATCT 787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 828 GGAAGTGTCTGACATGATGACACACAGAAATTCGACTTTGGAAGACTATCTCTG 887
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 788 TGAAGTATCTGACATGATGACACACAGAAATTTGACTTTGGAAGACTATCTCTG 847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 888 TCAGACACTGAATAACAACTTCTAAGCTGTTTACTAAGCTTAATCTCTGCAAAA 947
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 848 TCAGACACTGAATAACAACTTCTAAGCTGTTTACTAAGCTTAATCTCTGCAAAA 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 948 ATACAGTATTCCTATCATACTTGCAGACAGATTTGATGCTTGTCTCAACAGGTC 1007
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 908 GTACAGACTTCCATTTGATTAACAGAGATTTGATGCTTGTCTCAACAGGTC 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1008 TGGGAAGTCTGGGCTTTTCTCTACCAATTTGGCTCATATGATGATGATGAATAC 1067
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 968 TGGGAAGTCTGGGCTTTTCTCTCTATTTGGCTCATATGATGATGATGAATAC 1027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1068 TGCAGTCTGTTTAAAGAGTTGCAGAAACAGAGTATATTTAGCAACCACTGAGA 1127
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1028 TGCCAGTCTGTTTAAAGAGTTGCAGAAACAGAGTATATTTAGCAACCACTGAGA 1087
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1128 ATTGTCACACAGATTTATTTGGAAGCAGAAATTTCTTTTGGACCTTGTGAAGC 1187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1088 ACTGATCAACCAATTTACTTGGAGCCAGAAATTTCTTTTGGGCTTGTGTGAAGC 1147
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1188 TGTGTTATATATGGGGGAACCAAGCTGGGACATTCATTCGAACAATATGACAAGCTG 1247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1148 TGTGTTATATATGGGGGAACCAAGCTGGGACATTCATTCGAACAATATGACAAGCTG 1207
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1248 TAAATATATATGCTACTCTCGAAGACTGATGATATCATAGCAAGAAAGATTTG 1307
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1208 TAAATATATGCTACTCTCGAAGACTGATGATATCATAGCAAGAAAGATTTG 1267
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1308 TCTCAACAGATCAATTAATCTTATGTTGATGATGATGATGATGATGATGATG 1367
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1268 CTTCAACAGATCAATTAATCTTATGTTGATGATGATGATGATGATGATGATG 1327
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1368 TGTGTCAGAAATGAAGATTAATCTTGTCCAGGAATGCATCAAGGAACAGGCCA 1427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1328 TGAACCAAGATGAAGATTAATCTTGTCCAGGAATGCATCAAGGAACAGGCCA 1387
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1428 AACCTTATGTCAGTCAACTTTTCCAGAGAAATTCAAAGGTTGCTGACAGATTTT 1487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1388 AACCTTATGTCAGTCAACTTTTCCAGAGAAATTCAAAGGTTGCTGACAGATTTT 1447
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1488 AAAGTCAATTAATCTGTTGTTGCTGTTGGAACAAGTGGTGAAGATGATGATG 1547
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1448 AAAGTCAATTAATCTGTTGTTGCTGTTGGAACAAGTGGTGAAGATGATGATG 1507
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1548 GCAAGCCGTTCTCAAGTTGGCCAGTTCGAAAAGGAAGAACTGTTGAAATTCGCG 1607
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1508 GCAAGCCGTTCTCAAGTTGGCCAGTTCGAAAAGGAAGAACTGTTGAAATTCGCG 1567
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1608 AAACATAGGGAGTGAAGAACTATGCTTTGTTGAAACTTAAGAAAACAGATTTTAC 1667
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1568 AAACATAGGGAGTGAAGAACTATGCTTTGTTGAAACTTAAGAAAACAGATTTTAC 1627
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1668 TGCACTTTTCTTTGTCAGAAAAAATATCACTCAAGTATCATGCTGATGGGACA 1727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1628 TGCACTTTTCTTTGTCAGAAAAAATATCACTCAAGTATCATGCTGATGGGACA 1687
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1728 GAGAGACGGGAGCAAGCTTGTGAAGTTTCTGCTTGGAAAGTCCCAATTTCTGTTGC 1787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1688 GAGAGACGGGAGCAAGCTTGTGAAGTTTCTGCTTGGAAAGTCCCAATTTCTGTTGC 1747
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```

QY 1788 TACTTCAATGCTGCCAGAGGCTGATATTTGAAAAATGTGACATGTTTCAATTTTGA 1847
DB 1748 TACTTCAATGCTGCCAGAGGCTGATATTTGAAAAATGTGACATGTTTCAATTTTGA 1807
QY 1848 TCTTCCCTTACCAATGATATGTCATCGAATGGGGCTACGTGCTGTTGGGAA 1907
DB 1808 CTTTCTTCAACATGATATGTTTCAATGATGAGCAGCTGCGCTGTGAAA 1867
QY 1908 TACTGCGAGCAATTTCTTTTGTGATCTTGAATCGAATACCATTTAGACAGCTCT 1967
DB 1868 TACTGCGAGGCAATTTCTTTTGTGATCTGATCTGATATCATTTAGACAGCTCT 1927
QY 1968 AGTAAAGTATGACAGATGCTCAACAGATGTTCTGATGTTGGAAGAAATGCTCT 2027
DB 1928 AGTAAAGTATGACAGATGCTCAACAGATGTTCTGATGTTGGAAGAAATGCTCT 1987
QY 2028 TACTTCAATGCTGCCAGAGGCTGATATTTGAAAAATGTGACATGTTTCAATTTTGA 2081
DB 1988 TACTTCAATGCTGCCAGAGGCTGATATTTGAAAAATGTGACATGTTTCAATTTTGA 2047
QY 2082 AGTAAAGTATGACAGATGCTCAACAGATGTTCTGATGTTGGAAGAAATGCTCT 2132
DB 2048 TACTTCAATGCTGCCAGAGGCTGATATTTGAAAAATGTGACATGTTTCAATTTTGA 2107
QY 2133 TTTCAAGAGCTCCCAATCCAGTATGATGATGATGATGATGATGATGATGATGATGAT 2172
DB 2108 TTTCAAGAGCTCCCAATCCAGTATGATGATGATGATGATGATGATGATGATGATGAT 2147

```

```

RESULT 3
LOCUS BG717110 798 bp mRNA linear EST 08-MAY-2001
DEFINITION 60268919.F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821503 5',
            mRNA sequence.
ACCESSION BG717110
VERSION BG717110.1 GI:13996297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 798)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LHAM10728 row: e column: 24
            High quality sequence stop: 770.
            Location/Qualifiers
                1..798
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4821503"
                /clone_1db="NIH_MGC_97"
                /lab_host="DH10B"
                /note="Organ: testis; Vector: pBluescript (modified
                pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (9cgaag
                ) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVA-3',
                size-selected for average insert size 2.2 kb and
                normalized to ROP 5. This is a primary library enriched
                for full-length clones and constructed using the
                Cap-trapper method (Carninci, in preparation). Library
                constructed by M. Brownstein (NIH/NHGRI, National
                Institutes of Health). Note: this is a NIH_MGC Library."

```

## FEATURES

source

```

1..798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4821503"
/clone_1db="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (9cgaag
) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVA-3',
size-selected for average insert size 2.2 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

```

```

BASE COUNT 254 a 124 c 233 g 186 t 1 others
ORIGIN
Query Match 31.4%; Score 683; DB 12; Length 798;
Best Local Similarity 96.4%; Pred. No. 1.5e-181;
Matches 720; Conservative 0; Mismatches 25; Indels 2; Gaps 2;
QY 1 ATGGGGATGAAATGGGAGCAAAATCAACCTCATATGCTTCTCTATGTTCCATA 60
DB 52 ATGGAGATGAAAG- TGGGAGCAGAAATGACCCCTCATATGCTTCTCTATGTTCCATA 110
QY 61 TTTGAGAGATAGATATTTCTGAGAAATGAGACATTTTAAACAGAGCTCAGCTTCA 120
DB 111 TTTGAGAGATAGATATTTCTGAGAAATGAGACATTTTAAACAGAGCTCAGCTTCA 170
QY 121 TCATCAAAATGATATGATGAGACCTTCTCGAAGAGATCATTTCAATGAAATGATTTGCC 180
DB 171 TCATCAAAATGATATGATGAGACCTTCTCGAAGAGATCATTTCAATGAAATGATTTGCC 230
QY 181 TCTGGCGGAATTTTGGAAACAGAGATGCTGTGAGTATATAGCAGATTAATACATCC 240
DB 231 TCTGGCGGAATTTTGGAAACAGAGATGCTGTGAGTATATAGCAGATTAATACATCC 290
QY 241 ACAATGGGTGTTTGGAGTTGGAAGAGTTTGGAAACAGAGTTTTCAAACAGCAGG 300
DB 291 ACAATGGGTGTTTGGAGTTGGAAGAGTTTGGAAACAGAGTTTTCAAACAGCAGG 350
QY 301 TTTGAGAGATGATATGATGAGTCTGTTTCTGAGAGAGCTGATATGATGATGATGAT 360
DB 351 TTTGAGAGATGATATGATGAGTCTGTTTCTGAGAGAGCTGATATGATGATGATGAT 410
QY 361 CCAACACGGAACAGAGGTTTCCAAAGAGAGCGGCTATGAGATGGAATTAATTCAGAA 420
DB 411 CCAACACGGAACAGAGGTTTCCAAAGAGAGCGGCTATGAGATGGAATTAATTCAGAA 470
QY 421 GCTTCAGGGCCATACAGAGAGGTGGAAGAGTATTTCCGAGTTGCCGTGAGAGATT 480
DB 471 GCTTCAGGGCCATACAGAGAGGTGGAAGAGTATTTCCGAGTTGCCGTGAGAGATT 530
QY 481 GGTCTAGAGATGCCAATAATATGATTAAGACCCAGAGCAATGATGACGCCACTGTGGC 540
DB 531 GGTCTAGAGATGCCAATAATATGATTAAGACCCAGAGCAATGATGACGCCACTGTGGC 590
QY 541 CTTTGGGTCTTGAAGACAGCATTAATAGTGGACAGTATAGTGTACTTCTCAAGC 600
DB 591 CTTTGGGTCTTGAAGACAGCATTAATAGTGGACAGTATAGTGTACTTCTCTCAAGC 650
QY 601 AGAAGTGCAGTGAAGTGAACGAGGTGTTTACAAAGTTTAAATGAAGATTAATACA 660
DB 651 AGAAGTGCAGTGAAGTGAACGAGGTGTTTACAAAGTTTAAATGAAGATTAATACA 710
QY 661 GGTCTGGAAGAAATTTCTGGAAGTCAAGAACGAGAGAGAGAAAGT-AGTATATCTCA 719
DB 711 GGTCTGGAAGAAATTTCTGGAAGTCAAGAACGAGAGAGAGAAAGTATATCTCA 770
QY 720 AGGACCAAAAGTACCTACATACCCCC 746
DB 771 AGGACCAAAAGTACATTAATTTACCCC 797

```

```

RESULT 4
LOCUS B1464267 822 bp mRNA linear EST 21-AUG-2001
DEFINITION 603203379.F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269434 5',
            mRNA sequence.
ACCESSION B1464267
VERSION B1464267.1 GI:15254923
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 822)

```

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Miklos Palokvite, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NMGRI), Shiroki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Inyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM1679 row: m column: 19  
 High quality sequence stop: 714.  
 Location/Qualifiers

## FEATURES

## source

1..822  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5269434"  
 /clone\_lib="NIH MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescript (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
 ); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
 size selected for average insert size 2.2 kb and  
 normalized to 10<sup>5</sup>. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NMGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 255 a 140 c 239 g 188 t  
 ORIGIN

Query Match 29.5%; Score 640.8; DB 13; Length 822;  
 Best Local Similarity 95.4%; Pred. No. 1.3e-169;  
 Matches 725; Conservative 0; Mismatches 27; Indels 8; Gaps 6;  
 1 ATGGGGGATGAAGTTGGGAAGCAATCAACCTATATGCTTCTATGTTCCATA 60  
 64 ATGGGAAGTGAAGTTGGGAAGCAATCAACCTATATGCTTCTATGTTCCATA 123  
 61 TTGGAGAAGATAGTATTTCTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 120  
 124 TTGAGAAGATAGTATTTCTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 183  
 121 TCATCAGAAATGATGATGACCTTCTCGAAGATCATTTTCATGAAAAGTATTTGCC 180  
 184 TCATCAGAAATGATGATGACCTTCTCGAAGATCATTTTCATGAAAAGTATTTGCC 243  
 181 TTCTGGGGGATTTTGGAAAGAGATGCTGTGAGTAAATTAAGCAGATTAATCATCC 240  
 244 TTCTGGGGGATTTTGGAAAGAGATGCTGTGAGTAAATTAAGCAGATTAATCATCC 303  
 241 ACAATGGGTGTTTGGAGTTGAAAAGATTTTGAACACAGAGTTTTCACACAGAG 300  
 304 ACAATGGGTGTTTGGAGTTGAAAAGATTTTGAACACAGAGTTTTCACACAGAG 363  
 301 TTGGAAGTGTGATGATGCTGTGTTCTGAGAGAGTCTAGTAATGACTCGGAATAT 360  
 364 TTGGAAGTGTGATGATGCTGTGTTCTGAGAGAGTCTAGTAATGACTCGGAATAT 423  
 361 CCAACACGGAACAGAGGTTTTCACAGAGAGCGGCTATCGAGTGAATAATTCAGAA 420  
 424 CCAACACGGAACAGAGGTTTTCACAGAGAGCGGCTATCGAGTGAATAATTCAGAA 483  
 421 GCTTCAGGGCCATACAGAGAGGTGA-AGAGTAGTTTCCAGAGTTGCCGTGG-AGGAT 478  
 484 GCTTCAGGGCCATACAGAGAGGTGA-AGAGTAGTTTCCAGAGTTGCCGTGG-AGGAT 543  
 479 TTGGTCTAGGAAGTCCAAATATATGACTTGAACCCAGAGCAATGTATGACGGCACTGGTG 538  
 544 TTGGTCTAGGAAGTCCAAATATATGACTTGAACCCAGAGCAATGTATGACGGCACTGGTG 603

QY 539 GCCTTTTGGTCTAGAAAGACCAATTAAGTGCACAGATTAAGTACTTCTCAA 598  
 DB 604 GCC-TTTTGGTCTAGAAAGACCAATTAAGTGCACAGATTAAGTACTTCTCAA 662  
 QY 599 GCAGAAAGTGCAGTGAAGTGAAGCAAGGTGTTTACAAAGTTTAAATGAAGATATA 658  
 DB 663 GCAGAAAGTGCAGTGAAGTGAAGCAAGGTGTTTAAATGAAGATATA 722  
 QY 659 CAGG---CTCTGGAAGAATCTTGAAGTCAG-AAGCGAAGAGAGAGAAAGTAGAT 714  
 DB 723 CAGGCTTGGGCAAGAAATCTTGGGAAGTCAGAAAGCAAGCCGACAGAAAGTAGAT 782  
 QY 715 ACTCAAGG-ACCAAAAGTGACCTACATACCCCTCTCCA 753  
 DB 783 ACCCAAGAACCAAAAGTGACCTACATACCCCTCTCCA 822

RESULT 5  
 BQ429126 792 bp mRNA linear EST 24-MAY-2002  
 LOCUS AGENCOURT\_7901787 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6104663  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ429126  
 VERSION BQ429126.1 GI:21168202  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
 1 (bases 1 to 792)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM2342 row: n column: 24  
 High quality sequence stop: 496.  
 Location/Qualifiers

## FEATURES

## source

1..792  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6104663"  
 /clone\_lib="NIH MGC\_82"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgctcgcc); Site 2: SfiI (ggccatattggcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAGGCCATATGGCC-3' and 3' adaptor sequence:  
 5'-ATTCTGAGGCGCAGGCGGCCGATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo  
 Alto, CA)."  
 BASE COUNT 241 a 128 c 242 g 180 t 1 others  
 ORIGIN

Query Match 29.1%; Score 631.8; DB 14; Length 792;  
 Best Local Similarity 96.9%; Pred. No. 4.6e-167;  
 Matches 665; Conservative 0; Mismatches 18; Indels 3; Gaps 2;

QY 1 ATGGGGATGAAGTTGGGAAGCAATCAACCTATATGCTTCTATGTTCCATA 60  
 DB 61 ATGGGAATGAAGTTGGGAAGCAATCAACCTATATGCTTCTATGTTCCATA 120

QY 61 TTGAGAGAGATAGTATTTCTGAGAAATGAGACAATTTTAAACAGACTCCAGCTTCA 120  
 DB 121 TTGAGAGAGATAGTATTTCTGAGAAATGAGACAATTTTAAACAGACTCCAGCTTCA 180  
 QY 121 TCATCAGAAATGATGATGACCTTCTGAGAGATATTCATGAAAGATGATTTGGCC 180  
 DB 181 TCATCAGAAATGATGATGACCTTCTGAGAGATATTCATGAAAGATGATTTGGCC 240  
 QY 181 TCTGGCGGAAATTTGGAAACAGAGATGCTGATGATTAATGAGAGATATATACATCC 240  
 DB 241 TCTGGCGGAAATTTGGAAACAGAGATGCTGATGATTAATGAGAGATATATACATCC 300  
 QY 241 ACAATGGGTGTTTGGAGATGGAAGATTTTGGAAACAGAGATTTTCAACAGCAGG 300  
 DB 301 ACAATGGGTGTTTGGAGATGGAAGATTTTGGAAACAGAGATTTTCAACAGCAGG 360  
 QY 301 TTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 DB 361 TTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 QY 361 CCAACAGGAAACAGAGGATTTTCCAGAGAGGCGCTATGAGATGGAATATTTCAAGAA 420  
 DB 421 CCAACAGGAAACAGAGGATTTTCCAGAGAGGCGCTATGAGATGGAATATTTCAAGAA 480  
 QY 421 GCTTCAGGCGCAATCAGAGAGGTGGAAGAGTATTCGAGGTTCCGCGTGAAGATTT 480  
 DB 481 GCTTCAGGCGCAATCAGAGAGGTGGAAGAGTATTCGAGGTTCCGCGTGAAGATTT 540  
 QY 481 GGTCTAGAGAGTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 539  
 DB 541 GGTCTAGAGAGTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 QY 540 CCTTTTGGTCTGAGAGACAGATTAAGTGCAGACAGTATGATGATGATGATGATGATGAT 599  
 DB 601 CCTTTTGGTCTGAGAGACAGATTAAGTGCAGACAGTATGATGATGATGATGATGATGAT 660  
 QY 600 CAGAAGTGGAGTGAAGTGAAGCAGGTGTTTACAAAGGTTTAAATGAAGATTAATAC 659  
 DB 661 CAAAGGGCCAAATGAGATGGAACAGGGGTTTACAAAGGTTTAAATGAAGATTAATAC 720  
 QY 660 AGG--CTCTGGAAGAATTTCTTGGA 683  
 DB 721 CGGGCTCTGGAAGAATTTCTTGGA 746  
  
 RESULT 6  
 LOCUS BG719449 747 bp mRNA linear EST 08-MAY-2001  
 DEFINITION 602690172F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4822528 5',  
 mRNA sequence.  
 BG719449  
 ACCESSION BG719449  
 VERSION BG719449.1 GI:13998636  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 747)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabsd@imail.nih.gov  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LHAM10730 Row: P Column: 17  
 High quality sequence stop: 734.

FEATURES  
 source  
 Location/Qualifiers  
 1..747  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4822528"  
 /clone\_1db="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescriptPCr (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
 size-selected for average insert size 2.2 kb and  
 normalized to 10^5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 220 a 121 c 226 g 180 t  
 ORIGIN  
 Query Match 28.8%; Score 624.8; DB 12; Length 747;  
 Best Local Similarity 99.4%; Pred. No. 4.2e-165;  
 Matches 648; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
  
 QY 1 ATGGGGATGAAGATTTGGAGAGAGAAATCAACCTCATATGCTTCTATGTTCCATA 60  
 DB 97 ATGGAGATGAAGATTTGGAGAGAGAAATCAACCTCATATGCTTCTATGTTCCATA 156  
 QY 61 TTGAGAGAGATAGTATTTCTGAGAAATGAGACAATTTTAAACAGACTCCAGCTTCA 120  
 DB 157 TTGAGAGAGATAGTATTTCTGAGAAATGAGACAATTTTAAACAGACTCCAGCTTCA 216  
 QY 121 TCATCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 DB 217 TCATCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276  
 QY 181 TCTGGCGGAAATTTGGAAACAGAGATGCTGATGATTAATGAGAGATATATACATCC 240  
 DB 277 TCTGGCGGAAATTTGGAAACAGAGATGCTGATGATTAATGAGAGATATATACATCC 336  
 QY 241 ACAATGGGTGTTTGGAGATGGAAGATTTTGGAAACAGAGATTTTCAACAGCAGG 300  
 DB 337 ACAATGGGTGTTTGGAGATGGAAGATTTTGGAAACAGAGATTTTCAACAGCAGG 396  
 QY 301 TTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 DB 397 TTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456  
 QY 361 CCAACAGGAAACAGAGGATTTTCCAGAGAGGCGCTATGAGATGGAATATTTCAAGAA 419  
 DB 457 CCAACAGGAAACAGAGGATTTTCCAGAGAGGCGCTATGAGATGGAATATTTCAAGAA 516  
 QY 420 AGCTTCAGGCGCAATCAGAGAGGTGGAAGAGTATTCGAGGTTCCGCGTGAAGATTT 479  
 DB 517 AGCTTCAGGCGCAATCAGAGAGGTGGAAGAGTATTCGAGGTTCCGCGTGAAGATTT 576  
 QY 480 TGGTCTGAGAGTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539  
 DB 577 TGGTCTGAGAGTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 636  
 QY 540 CCTTTTGGTCTGAGAGACAGATTAAGTGCAGACAGTATGATGATGATGATGATGATGAT 599  
 DB 637 CCTTTTGGTCTGAGAGACAGATTAAGTGCAGACAGTATGATGATGATGATGATGATGAT 696  
 QY 600 CAGAAGTGGAGTGAAGTGAAGCAGGTGTTTACAAAGGTTTAAATGAAGAA 651  
 DB 697 CAGAAGTGGAGTGAAGTGAAGCAGGTGTTTACAAAGGTTTAAATGAAGAA 747  
  
 RESULT 7  
 LOCUS BG718234 682 bp mRNA linear EST 08-MAY-2001  
 DEFINITION 602696209F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4828298 5',  
 mRNA sequence.

ACCESSION BG18234  
 VERSION BG18234.1 GI:13997421  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 682)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 UNPUBLISHED (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LAM10746 row: a column: 03  
 High quality sequence stop: 682.  
 Location/Qualifiers  
 1. 682  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4828298"  
 /clone\_1lb="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag ) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 205 a 109 c 204 g 164 t  
 ORIGIN  
 Query Match 28.6%; Score 620.4; DB 12; Length 682;  
 Best Local Similarity 99.8%; Pred. No. 7.1e-164;  
 Matches 621; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGGGGGATGAAGTGGGAAGCAGAAATCAACCTCTATAGTCTTCTATGTTCCATA 60  
 DB 61 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCTCTATAGTCTTCTATGTTCCATA 120  
 QY 61 TTTGGAAGATGAGTATTCTGGAGAAATGAGACATTTTAAACAGACTCCAGCTTCA 120  
 DB 121 TTTGGAAGATGAGTATTCTGGAGAAATGAGACATTTTAAACAGACTCCAGCTTCA 180  
 QY 121 TCATCAGAAATGATGATGAGACCTTCTGGAAGAGATCATTTTCATGAAAAAGTGAATGGC 180  
 DB 181 TCATCAGAAATGATGATGAGACCTTCTGGAAGAGATCATTTTCATGAAAAAGTGAATGGC 240  
 QY 181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGCTGAGTAAATTAAGCGATTAATACATCC 240  
 DB 241 TCTGGGCGGAATTTTGGAAACAGAGATGCTGCTGAGTAAATTAAGCGATTAATACATCC 300  
 QY 241 ACAATGGGTGTTTGGAGTTGGAAGATGTTTGGAAACAGAGTTTTCACACGACG 300  
 DB 301 ACAATGGGTGTTTGGAGTTGGAAGATGTTTGGAAACAGAGTTTTCACACGACG 360  
 QY 301 TTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 DB 361 TTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 QY 361 CCAACACGGAACAGAGGGTTTCCAGAGAGCGGCTATGAGATGAGAAATTAATTCAGAA 420  
 DB 421 CCAACACGGAACAGAGGGTTTCCAGAGAGCGGCTATGAGATGAGAAATTAATTCAGAA 480

QY 421 GCTTCAGGGCCATACAGAAAGGTGGAAGATGATTTCCAGAGTTGCCGTGAGATTT 480  
 DB 481 GCTTCAGGGCCATACAGAAAGGTGGAAGATGATTTCCAGAGTTGCCGTGAGATTT 540  
 QY 481 GGTCTAGGAAGTCCAAATAATGATGATGATGATGATGATGATGATGATGATGAT 540  
 DB 541 GGTCTAGGAAGTCCAAATAATGATGATGATGATGATGATGATGATGATGATGAT 600  
 QY 541 CTTTGGTCTTGAAGACCAAGTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 600  
 DB 601 CTTTGGTCTTGAAGACCAAGTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 660  
 QY 601 AGAAGTGCAGTGAAGTGAAC 622  
 DB 661 AGAAGTGCAGTGAAGTGAAC 682  
 RESULT 8  
 B1560849  
 LOCUS  
 DEFINITION B1560849 693 bp mRNA linear EST 05-SEP-2001  
 603254018F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5296444 5',  
 mRNA sequence.  
 ACCESSION B1560849  
 VERSION B1560849.1 GI:15448163  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 693)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 UNPUBLISHED (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LAM11750 row: c column: 05  
 High quality sequence stop: 691.  
 Location/Qualifiers  
 1. 693  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5296444"  
 /clone\_1lb="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag ) ; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 207 a 111 c 207 g 168 t  
 ORIGIN  
 Query Match 28.6%; Score 620.4; DB 13; Length 693;  
 Best Local Similarity 99.8%; Pred. No. 7.1e-164;  
 Matches 621; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGGGGGATGAAGTGGGAAGCAGAAATCAACCTCTATAGTCTTCTATGTTCCATA 60  
 DB 72 ATGGGAGATGAAGATTGGGAAGCAGAAATCAACCTCTATAGTCTTCTATGTTCCATA 131

QY	61	TTTGAGAAAGATAGGATATCTCGAGAAATGAGACATTTTAAAGAGACTCCACTTCA	120
Db	132	TTTGAGAAAGATAGGATATCTCGAGAAATGAGACATTTTAAAGAGACTCCACTTCA	191
QY	121	TCATCAGAAATGATGATGACCTTCTCGAAGAGATCATTTTCATGAAAGTGGATTGGC	180
Db	192	TCATCAGAAATGATGATGACCTTCTCGAAGAGATCATTTTCATGAAAGTGGATTGGC	251
QY	181	TCGTGGCCGGAATTTTGGAAAACAGAGATCGTGTGATGCTAATAGCGAGATTAATCATCC	240
Db	252	TCGTGGCCGGAATTTTGGAAAACAGAGATCGTGTGATGCTAATAGCGAGATTAATCATCC	311
QY	241	ACAATGGGTGGTTTGGAGTTGGAAAAGTTTGGAAAACAGAGCTTTTCAAAACGACAGG	300
Db	312	ACAATGGGTGGTTTGGAGTTGGAAAAGTTTGGAAAACAGAGCTTTTCAAAACGACAGG	371
QY	301	TTTGAAGATGGTGATAGCTCTGTTCTTCGAGAGAGCTAGTAATAGCTGCGAGATAT	360
Db	372	TTTGAAGATGGTGATAGCTCTGTTCTTCGAGAGAGCTAGTAATAGCTGCGAGATAT	431
QY	361	CCAACACGGAAACAGAGGCTTTTCCAAAGAGAGCGGCTTCCGAGATGAAATTAATTCAGAA	420
Db	432	CCAACACGGAAACAGAGGCTTTTCCAAAGAGAGCGGCTTCCGAGATGAAATTAATTCAGAA	491
QY	421	GCTTCAGGCGCATACAGAAAGGTGGAGAGGTAGTTTCCGAGGTTCGCGTGGAGAGATT	480
Db	492	GCTTCAGGCGCATACAGAAAGGTGGAGAGGTAGTTTCCGAGGTTCGCGTGGAGAGATT	551
QY	481	GGCTAGGAATCCAAATATATGACTTTAGACCACAGATGATGATCAGCGCACTGTGGC	540
Db	552	GGCTAGGAATCCAAATATATGACTTTAGACCACAGATGATGATCAGCGCACTGTGGC	611
QY	541	CTTTTGGTCTAGAAACCCAGTATTAAGTGGCACAGGTATGTGATACCTTCCAAAGC	600
Db	612	CTTTTGGTCTAGAAACCCAGTATTAAGTGGCACAGGTATGTGATACCTTCCAAAGC	671
QY	601	AGAAGTGCAGTGGAAATGAAAC	622
Db	672	AGAAGTGCAGTGGAAATGAAAC	693
RESULT 9			
BI462207			
LOCUS			
DEFINITION			
BI462207 830 bp mRNA linear EST 21-AUG-2001			
6032055201 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5271221 5',			
mRNA sequence.			
ACCESSION			
BI462207			
VERSION			
BI462207.1 GI:15252863			
KEYWORDS			
EST.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE			
1 (bases 1 to 830)			
NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
AUTHORS			
National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE			
Unpublished (1999)			
JOURNAL			
COMMENT			
Contact: Robert Strausberg, Ph.D.			

```

FEATURES
  source      location/Qualifiers
              1. .830
              /organism="Homo sapiens"
              /db_xref="taxon:9606"

```

	/clone="IMAGE:5271221"	
	/clone_lib="NIH_MGC_97"	
	/lab_host="DHI0B"	
	/note="Organ: testis; Vector: pBluescriptPLR (modified	
	pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gscag	
	y); Oligo-dT primed using primer 5'-TTTTTTTATTTTATTTVN-3'	
	; size-selected for average insert size 2.2 kb and	
	normalized to ROT 5. This is a primary library enriched	
	for full-length clones and constructed using the	
	Cap-trapped method (Carninci), in preparation). Library	
	constructed by M. Brownstein (NIMH/NHGRI, National	
	Institutes of Health). Note: this is a NIH_MGC Library."	
BASE COUNT	255 a 137 c 244 g 193 t	1 others
ORIGIN		
Query Match	28.3%; Score 615.6; DB 13; Length 830;	
Best Local Similarity	95.8%; Pred. No. 1.8e-162;	
Matches 729; Conservative	0; Mismatches 20; Indels 12; Gaps 9;	
Oy	1 ATGGGSGATGAAGATTGGGAAGCAGAATAACACCCTCATATGTCTTCATATGCCATA	60
Db	72 ATGGGAAGATGAAGATTGGGAAGCAGAATAACACCCTCATATGTCTTCATATGCCATA	131
Oy	61 TTTGAGAAGTAGATTAATCTCGAAGAAATGGAGACAATTTAACAGAGACTCCAGCTCA	120
Db	132 TTTGAGAAGTAGATTAATCTCGAAGAAATGGAGACAATTTAACAGAGACTCCAGCTCA	191
Oy	121 TCATCAGAAATGATGATGACCTTCTCGAAGATCATTTTCATGAAAAGTGATTTGCC	180
Db	192 TCATCAGAAATGATGATGACCTTCTCGAAGATCATTTTCATGAAAAGTGATTTGCC	251
Oy	181 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGGTAGTGAATTAAGCAGATAAATCATCC	240
Db	252 TCTGGGCGGAATTTTGGAAAACAGAGATGCTGGTAGTGAATTAAGCAGATAAATCATCC	311
Oy	241 ACAATGGGTGTTTGGAGTTGGAAAGATTTTGGAAACAGAGTTTTCACACAGAG	300
Db	312 ACAATGGGTGTTTGGAGTTGGAAAGATTTTGGAAACAGAGTTTTCACACAGAG	371
Oy	301 TTGGAAGATGGATPACTCTGATTTTTCGAGAGAGCTGATATGACTGCGAAGATPAT	360
Db	372 TTGGAAGATGGATPACTCTGATTTTTCGAGAGAGCTGATATGACTGCGAAGATPAT	431
Oy	361 CCAAACAGGAACAGAGGGTTTCCAAAGAGAGCGGCTATCGAGATGGAATAATTCAGAA	420
Db	432 CCAAACAGGAACAGAGGGTTTCCAAAGAGAGCGGCTATCGAGATGGAATAATTCAGAA	491
Oy	421 GCTTCAGGGCATPACAGAAAGGTGGAAAGATAGTTTCCGAGGTTCCTGGAGAGATT	480
Db	492 GCTTCAGGGCATPACAGAAAGGTGGAAAGATAGTTTCCGAGGTTCCTGGAGAGATT	550
Oy	481 GGTCAGGAAGTCCAAATTAATGATTATGACCCACAGCAATGTATGACAGCGCACGTGGGC	540
Db	551 GGTCAGGAAGTCCAAATTAATGATTATGACCCACAGCAATGTATGACAGCGCACGTGGGC	610
Oy	541 CTTTTGGTTCTAGAAAGCAGTATTAAGT-GGCACAGTAAATGTGATATCTTCTCAAG	599
Db	611 C-TTTTGGTTCTAGAAAGCAGTATTAAGTGGGCACAGTAAATGTGATATCTTCTCAAG	669
Oy	600 CAGAAATGGCAG-TGGAAGTGAAGC-AGGTGGTTACAAAGTTT--AAATGAAGAACTA	654
Db	670 CAGAAATGGCAGTGGTGAAGTGAAGCAAGGTGGTTTACAAAGTTTCACTGMAAGAACTA	729
Oy	655 ATAACAGAGCTCTGAA-AGAATTTTGGAAAGTCAAG-GCAGAAAGAGGA-GAAAGTGA	710
Db	730 ATAACAGAGGTTGGAACAGACTTTTGGAAAGTCAAGAAAGGAGAGAGCAACAGTAG	789
Oy	711 TGATATCTCAAGSACCAAAAGTGAACCTACATACCCCTCTCTC 751	
Db	790 TGATATCTCAAGSACCAAAAGTGAACCTAATATCCCCGCTCTTC 830	

B1561001  
 LOCUS 918 bp mRNA linear EST 05-SEP-2001  
 DEFINITION 603254391F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:529534 5',  
 mRNA sequence.  
 ACCESSION B1561001  
 VERSION B1561001.1 GI:15448315  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 918)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9abs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LHAM1750 row: f column: 23  
 High quality sequence stop: 745.  
 Location/Qualifiers  
 source  
 1. 918  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:529534"  
 /clone\_1db="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescript (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gucgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',  
 size selected for average insert size 2.2 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 288 a 160 c 253 g 217 t  
 ORIGIN  
 Query Match 28.2%; Score 612.4; DB 13; Length 918;  
 Best Local Similarity 86.8%; Pred. No. 1.5e-161;  
 Matches 753; Conservative 0; Mismatches 36; Indels 79; Gaps 4;  
 QY 1 ATGGGGGATGAAGTTTGGGAAGCAATCAACCTCATATGCTTCTTATGTTCCATA 60  
 DB 61 ATGGGAATGAAGTTTGGGAAGCAATCAACCTCATATGCTTCTTATGTTCCATA 120  
 QY 61 TTTGAGAAGATAGTATTTCTGAGAAATGAGACATTTTAAACAGACTCCAGTTTCA 120  
 DB 121 TTTGAGAAGATAGTATTTCTGAGAAATGAGACATTTTAAACAGACTCCAGTTTCA 180  
 QY 121 TCATCAGAAATGATGATGACCTTCTCGAAGAGATCATTTTCATGAAGAGATTGGCC 180  
 DB 181 TCATCAGAAATGATGATGACCTTCTCGAAGAGATCATTTTCATGAAGAGATTGGCC 240  
 QY 181 TCTGGGGGGAATTTTGGAAACAGAGATGCTGCTAGTAAAGAGATTAACATCC 240  
 DB 241 TCTGGGGGGAATTTTGGAAACAGAGATGCTGCTAGTAAAGAGATTAACATCC 270  
 QY 241 ACAATGATGATTTTGGAGTTGAAAGATTTTGGAAACAGAGATTTTCAACAGAG 300  
 DB 271 -----GGTTTTTCAACAGAGG 288  
 QY 301 TTTGAAGATGATGATGATGATGATTTTCTGAGAGAGATCTAGTAATGACTGGAGATTAAT 360  
 DB 289 TTTGAAGATGATGATGATGATGATTTTCTGAGAGAGATCTAGTAATGACTGGAGATTAAT 348

QY 361 CCAACACGAGACAGAGGGTTTCCAAAGAGAGCGGCTATCGAGATGAAATATTTCAGAA 420  
 DB 349 CCAACACGAGACAGAGGGTTTCCAAAGAGAGCGGCTATCGAGATGAAATATTTCAGAA 408  
 QY 421 GCTTCAGGGCCATACAGAAAGAGTGAAGAGGTTTCCAGAGTTGCCGTGAGAGATT 480  
 DB 409 GCTTCAGGGCCATACAGAAAGAGTGAAGAGGTTTCCAGAGTTGCCGTGAGAGATT 468  
 QY 481 GGTCTGGAAGTCCAAATTAATGACTTAAGCCAGACGAATGATGAGGCACTGGTGC 540  
 DB 469 GGTCTGGAAGTCCAAATTAATGACTTAAGCCAGACGAATGATGAGGCACTGGTGC 528  
 QY 541 CTTTGGTTCTTGAAGAACCAATTAATGAGTGGACACAGTAATGATGATCTTCTCAAC 600  
 DB 529 CTTTGGTTCTTGAAGAACCAATTAATGAGTGGACACAGTAATGATGATCTTCTCAAC 588  
 QY 601 AGAAGTGCAGTGAAGTGAACGAGGTGTTCAAAAGTTTAAATGACAGATTAATACA 660  
 DB 589 AGAAGTGCAGTGAAGTGAACGAGGTGTTCAAAAGTTTAAATGACAGATTAATACA 648  
 QY 661 GGCTCTGGAAGAAATTTCTTGAAGTCAAGAGAGAGAGAGAGAG- AAAGTATGATCTCA 719  
 DB 649 GGCTCTGGAAGAAATTTCTTGAAGTCAAGAGAGAGAGAGAGAGAGATGATCTCA 708  
 QY 720 AGACACCAAGTGAACCTTACATACCCCTCTCCACCTGAGATGAGG- ACTTCATC 774  
 DB 709 AGACACCAAGTGAACCTTACATACCCCTCTCCACCTGAGATGAGGACTTCATCTT 768  
 QY 775 TTTGACATTTATCAGACAGCATTAAC-TTGACAAATTCAGACACTATTCTTGTGAGT 833  
 DB 769 TTTGACATTTATCAGACAGCATTAAC-TTGACAAATTCAGACACTATTCTTGTGAGT 828  
 QY 834 GTCTGACATGATGACACCAACGACATTT 861  
 DB 829 GTCTGACATTTATGACACCAACGACATTT 856  
 RESULT 11  
 LOCUS BG718634 846 bp mRNA linear EST 08-MAY-2001  
 DEFINITION 602696739F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4828977 5',  
 mRNA sequence.  
 ACCESSION BG718634  
 VERSION BG718634.1 GI:13997821  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 846)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9abs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LHAM10747 row: m column: 10  
 High quality sequence stop: 701.  
 Location/Qualifiers  
 source  
 1. 846  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4828977"  
 /clone\_1db="NIH\_MGC\_97"  
 /lab\_host="DH10B"

/note="Organ: testis; Vector: pbluescriptR (modified pbluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGR), National Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 251 a 153 c 248 g 194 t

ORIGIN

Query Match 28.1%; Score 610.4; DB 12; Length 846;  
Best Local Similarity 94.7%; Pred. No. 5.3e-161;  
Matches 720; Conservative 0; Mismatches 26; Indels 14; Gaps 8;

QY 1 ATGGGGGATGAATTTGGGAGGAAATGACCTCATATGCTTCTTCTTCTTCCATA 60  
DB 72 ATGGGAGATGAATTTGGGAGGAAATGACCTCATATGCTTCTTCTTCCATA 131  
QY 61 TTGAGAGAGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 120  
DB 132 TTGAGAGAGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 131  
QY 121 TCATCAGAAATGATGATGAGACCTTCTCGAAGAGATCATTTGAAAAGTGGATTGCC 180  
DB 192 TCATCAGAAATGATGATGAGACCTTCTCGAAGAGATCATTTGAAAAGTGGATTGCC 251  
QY 181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGTGAGTATTAAGCAGATTAATCATCC 240  
DB 252 TCTGGGCGGAATTTTGGAAACAGAGATGCTGTGAGTATTAAGCAGATTAATCATCC 311  
QY 241 ACAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300  
DB 312 ACAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 371  
QY 301 TTGAGAGATGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 360  
DB 372 TTGAGAGATGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 431  
QY 361 CCAACACGGAACGAGAGGTTTCCAGAGAGCGGCTATGAGATGGAATTAATTCAGAA 420  
DB 432 CCAACACGGAACGAGAGGTTTCCAGAGAGCGGCTATGAGATGGAATTAATTCAGAA 491  
QY 421 GCTTCAGGCGCATACAGAGAGGTTG- AAGAGTATGTTCCGAGTGGCGGCTGG-AGGAT 478  
DB 492 GCTTCAGGCGCATACAGAGAGGTTG- AAGAGTATGTTCCGAGTGGCGGCTGG-AGGAT 551  
QY 479 TTGGTCTAGGAAGTCCAAATTAATGACTTAGACCCAGACGAATGATGACAGCACTGGTG 538  
DB 552 TTGGTCTAGGAAGTCCAAATTAATGACTTAGACCCAGACGAATGATGACAGCACTGGTG 611  
QY 539 GCTTTTGGTCTTAGAAGCCAGTATT- AAGTGGCAGAGTATGCTGATCTTCTCAA 597  
DB 612 GCTTTTGGTCTTAGAAGCCAGTATT- AAGTGGCAGAGTATGCTGATCTTCTCAA 671  
QY 598 AGCAGAGTGGCAGTGGAGTGAACGAGTGGTTACA- AAGTGGTAAATGAAGTAAT 656  
DB 672 AGCAGAGTGGCAGTGGAGTGAACGAGTGGTTACA- AAGTGGTAAATGAAGTAAT 731  
QY 657 AA-CAGGCTCTGGAAG--AATTCTTGAAGTCAGAGAGAGAGAGAG--AAAGT 708  
DB 732 AACCAAGGCTCTGGAAGCAATTTCTTGAACAGTCAAGAAAGCGGAGCGGAGCAACCGT 791  
QY 709 AGTGAATCTCAAGAGCAAAAGTACCTACATACCCCTTC 748  
DB 792 AGCGATACTCA--GGAACACAGTACCTACATACCCGCTC 829

RESULT 12  
LOCUS BG771822 699 bp mRNA linear EST 15-MAY-2001  
DEFINITION 602720504F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4837296 5',

ACCESSION mRNA sequence.  
VERSION BG771822  
KEYWORDS BG771822.1 GI:14082475  
SOURCE EST.  
ORGANISM human.

REFERENCE Homo sapiens  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
JOURNAL Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
COMMENT NIH-MGC http://imgc.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGR), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
http://image.llnl.gov  
plate: LLAM10769 row: h column: 01  
High quality sequence stop: 692.  
Location/Qualifiers  
1. 699  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1ib="NIH\_MGC\_97"  
/lab\_host="DH10B"

FEATURES  
source

/note="Organ: testis; Vector: pbluescriptR (modified pbluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGR), National Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 208 a 113 c 211 g 167 t

ORIGIN

Query Match 27.5%; Score 597.4; DB 12; Length 699;  
Best Local Similarity 99.5%; Pred. No. 2.3e-157;  
Matches 620; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 ATGGGGATGAAGATTGGGAGGAAATCAACCTCATATGCTTCTTCTTCCATA 60  
DB 78 ATGGGAGATGAAGATTGGGAGGAAATCAACCTCATATGCTTCTTCTTCCATA 137  
QY 61 TTGAGAGAGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 120  
DB 138 TTGAGAGAGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 197  
QY 121 TCATCAGAAATGATGATGAGACCTTCTCGAAGAGATCATTTGAAAAGTGGATTGCC 180  
DB 198 TCATCAGAAATGATGATGAGACCTTCTCGAAGAGATCATTTGAAAAGTGGATTGCC 257  
QY 181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGTGAGTATTAAGCAGATTAATCATCC 240  
DB 258 TCTGGGCGGAATTTTGGAAACAGAGATGCTGTGAGTATTAAGCAGATTAATCATCC 317  
QY 241 ACAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300  
DB 318 ACAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 376  
QY 301 TTGAGAGATGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 360  
DB 377 TTGAGAGATGATAGTATCTTGAGAAATGAGACAAATTTTAAACAGACTCCAGCTTCA 436  
QY 361 CCAACACGGAACGAGAGGTTTCCAGAGAGCGGCTATGAGATGGAATTAATTCAGAA 420

Db	CCAAACGGAACAGAGGGTTTCCAAAGAGAGCGCGTATCGAGATGGAATTAATTCAGAA	496
Qy	437	
Qy	421	GCTTCAGGGCCATACAGAAAGGTGTGAAGAGTACTGTTCCAGAGTTCCTGTGAGGATTT
Db	497	GCTTCAGGGCCATACAGAAAGGTGTGAAGAGTACTGTTCCAGAGTTCCTGTGAGGATTT
Qy	481	GGTCTAGGAATCCAAATTAATGACTTGTAGACCCAGACGGAATGTATGCAAGCCGACATGAGTGC
Db	557	GGTCTAGGAATCCAAATTAATGACTTGTAGACCCAGACGGAATGTATGCAAGCCGACATGAGTGC
Qy	541	CTTTTGTGTTCTAGNAGACAGTATTAAGTGGCAGAGTAAATGCTGATCTTCCAAGC
Db	616	CTTTTGTGTTCTAGNAGACAGTATTAAGTGGCAGAGTAAATGCTGATCTTCCAAGC
Qy	601	AGAACTGGCAGCTGAACTGAACG 623
Db	676	AGAACTGGCAGCTGAACTGAACG 698
RESULT 13		
LOCUS	BI559479	758 bp mRNA linear EST 05-SEP-2001
DEFINITION	603252932P1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5295407 5',	
ACCESSION	BI559479	
VERSION	BI559479.1	GI:15446793
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 758)	
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.	
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)	
	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.	
	Email: cgapbs-remail.nih.gov	
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.	
	cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki	
	Tochiyuki and Piero Carninci (RIKEN)	
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN).	
	DNA Sequencing by: Incyte Genomics, Inc.	
	Clone distribution: MGC clone distribution information can be	
	found through the I.M.A.G.E. Consortium/LNLN at:	
	http://image.llnl.gov	
	Plate: LLM11747 row: 9 column: 24	
FEATURES	High quality sequence stop: 741.	
source	location/Qualifiers	
	1..758	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5295407"	
	/clone_11b="NIH_MGC_97"	
	/lab_host="DH10B"	
	/note="Organ: testis; Vector: pBluescriptR (modified	
	pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag	
	); Oligo-dT primed using primer 5'-gtttttttttttttn-3',	
	size-selected for average insert size 2.2 kb and	
	normalized to ROT 5. This is a primary library enriched	
	for full-length clones and constructed using the	
	Cap-trapper method (Carninci, in preparation). Library	
	constructed by M. Brownstein (NIH/NHGRI, National	
	Institutes of Health). Note: This is a NIH_MGC Library."	
BASE COUNT	232 a 127 c 220 g 179 t	
ORIGIN		
Query Match	27.5%; Score 556.6; DB 13; Length 758;	
Best Local Similarity	91.2%; Pred. NO. 4e-157;	
Matches	681; Conservative 0; Mismatches 4; Indels 62; Gaps 2;	
Qy	1 ATGGGGATGTAAGATTGGAGGACAAATCAACCTCATATGTCCTCTATGTTCCCAATA	60
Db	62 ATGGGAGATGTAAGATTGGAGGACAAATCAACCTCATATGTCCTCTATGTTCCCAATA	121

Accession	Source	Organism	Version	Keywords	Reference	Title	Journal	Comment
61	TTTGAAGAAGTATGCTTTCTCGGAGAAATGGAGACATTTTAA	CAGAGATCTCCAGCTTCA	120					
122	TTTGAAGAAGTATGCTTTCTCGGAGAAATGGAGACATTTTAA	CAGAGATCTCCAGCTTCA	181					
121	TCATCAGAAATGATGATGACCTTCTCGAAGATCATTTATGAA	AGTGTGATTTTCC	180					
182	TCATCAGAAATGATGATGACCTTCTCGAAGATCATTTATGAA	AGTGTGATTTTCC	241					
181	TTCTGGGGGGAATTTTGGAAACAGAGATGCTGTGATTAAT	AGCAAGATTAATATCC	240					
242	TTCTGGGGGGAATTTTGGAAACAGAGATGCTGTGATTAAT	AGCAAGATTAATATCC	301					
241	ACAATGGGTGTTTGGAGTTGGAAGATTTTGGAAACAGAG	ATTTTCAACAGCAGG	300					
302	ACAATGGGTGTTTGGAGTTGGAAGATTTTGGAAACAGAG	ATTTTCAACAGCAGG	361					
301	TTTGAAGATGATGATGCTCTGTGTTCTGGAGACAGTCT	AGTAAATGACTCGGAAGATAT	360					
362	TTTGAAGATGATGATGCTCTGTGTTCTGGAGACAGTCT	AGTAAATGACTCGGAAGATAT	394					
361	CCAAACAGGAACAGAGGTTTCCAAAGAGCGCGCTATG	AGATGGAATTAATTCAGAA	420					
395	-----	-----	421					
421	GCTTCAGAGGCATACAGAGAGTGGAGAGATAGTTCCG	AGGTTCCCGTGGAGATTT	480					
422	GCTTCAGAGGCATACAGAGAGTGGAGAGATAGTTCCG	AGGTTCCCGTGGAGATTT	481					
481	GGTCTAGGAAGTCCAAATTAATGACTTAAGACCCAG	ACGAAATGTATGCAGCCGACATGCTGGCC	540					
482	GGTCTAGGAAGTCCAAATTAATGACTTAAGACCCAG	ACGAAATGTATGCAGCCGACATGCTGGCC	541					
541	CTTTTGGTCTTAAGAACCAGATTTAAGTGGGACAG	GTAAATGGTGAATCTTCCAAGC	600					
542	CTTTTGGTCTTAAGAACCAGATTTAAGTGGGACAG	GTAAATGGTGAATCTTCCAAGC	601					
601	AGAAATGGCAGTGAAGTGAACGAGTGTGTTACAA	AGTTTAAATGAA--GAAATTAAT	658					
602	AGAAATGGCAGTGAAGTGAACGAGTGTGTTACAA	AGTTTAAATGAA--GAAATTAAT	661					
659	CAGGCTCTGGAAGAAATTTCTTGAAGTCAAGAC	GAGAAAGGAGAAAGTATGATATCT	718					
662	CAGGCTCTGGAAGAAATTTCTTGAAGTCAAGAC	GAGAAAGGAGAAAGTATGATATCT	721					
719	AAGGACCAAAAGTACCTACATACCCC	745						
722	CAGGACCAAAAGTACCTACATACCCC	748						
RESULT 14								
BI461393	861 bp	mRNA	linear	EST 21-AUG-2001				
LOCUS	60320686F1 NIH_MGC_97 Homo sapiens	cdna	clone IMAGE:5272430 5'					
DEFINITION	mRNA sequence.							
ACCESSION	BI461393							
VERSION	BI461393.1	GI:15252049						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
	1 (bases 1 to 861)							
	NIH-MGC <a href="http://imgc.ncl.nih.gov/">http://imgc.ncl.nih.gov/</a> .							
	National Institutes of Health, Mammalian Gene Collection (MGC)							
	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>							
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.							
	cdna Library Preparation: Michael J. Brownstein (NHGRI), Shiroki							
	Toshiyuki and Piero Carninci (RIKEN)							
	cdna Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)							
	DNA Sequencing by: Incyte Genomics, Inc.							



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM1687 row: j column: 15  
 High quality sequence stop: 672.

FEATURES  
SOURCE

### SOURCE

```
//organism="Homo sapiens"  
//db_xref="taxon:9606"  
//clone="IMAGE:5872430"  
//clone_lib="NHT MGC_97"  
//lab_host="DH10B"  
  
//note= "Organ. testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gcgcga  
cggcgccgacggaggcttcccagatgttcgccccgcggcgagcaaa)  
"; Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',  
size selected for average insert size 2.2 kb and  
normalized to ROF 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI National  
Institutes of Health). Note: this is a NIH MGC Library."
```

	a	c	g	t	-
BASE COUNT	265	147	247	201	-
ORIGIN					

Query Match 27.4%; Score 594.8; DB 13; Length 861;

Best Local Similarity 89.7%; Pred. No. 1.4e-156;  
Matches 722; Conservative 0; Mismatches 63; Indels 20; Gaps 7;

QY	1	ATGGGGGATGAAGATTGGGAGCGAAGAAATCAACCTCATATGCTCTCTATGTTCCCATTA	60
Db	49	ATGGGAAGTGAAGATTGGAGCGCAAGAAATCAACCTCATATGCTCTCTATGTTCCCATTA	108
QY	61	TTTGAGAAAGATAGGTATTTCTGGAGAAAATGAGACAAATTTTAAACAGACTCCAGCTTCA	120
Db	109	TTTGAGAAAGATAGGTATTTCTGGAGAAAATGAGACAAATTTTAAACAGACTCCAGCTTCA	168
QY	121	TCATCAGAAAATGATGATGACCTTCCGAGAGATCATTTTCATGAAAAGTGGATTGGCC	180
Db	169	TCATCAGAAAATGATGATGACCTTCCGAGAGATCATTTTCATGAAAAGTGGATTGGCC	228
QY	181	TCCTGGGGCGAATTTTGGAAAACAGAGATGCTGGTAGTATTAATGAACGAGATTAATCATCC	240
Db	229	TCCTGGGGCGAATTTTGGAAAACAGAGATGCTGGTAGTATTAATGAACGAGATTAATCATCC	288
QY	241	ACAAATGGGTGTTTGGAGTTGGAAAGAGTTTGGAAAACAGAGTTTTCACAAACAGCAGG	300
Db	289	ACAAATGGGTGTTTGGAGTTGGAAAGAGTTTGGAAAACAGAGTTTTCACAAACAGCAGG	348
QY	301	TTTGAAGATGGTATGAGTCTGTGTTTCTGGAGAGAGTATAGTAATGAATCGGAGGAATAT	360
Db	349	TTTGAAGATGGTATGAGTCTGTGTTTCTGGAGAGAGTATAGTAATGAATCGGAGGAATAT	408
QY	361	CCAAACACGGAAACAGAGGGTTTCCAAAGAGAGCGGCTATGAGATGAAAATTAATTCAGAA	420
Db	409	CCAAACACGGAAACAGAGGGTTTCCAAAGAGAGCGGCTATGAGATGAAAATTAATTCAGAA	468
QY	421	GCTTCAGGGCCATACAGAGAGGGTGGAAAGAGTGTGTTCCAGAGTGGCCGGAGGAGATT	480
Db	469	GCTTCAGGGCCATACAGAGAGGGTGGAAAGAGTGTGTTCCAGAGTGGCCGGAGGAGATT	527
QY	481	GGTCTAGGAAGTCCAAATTAATGACTTTAGACCACAGACGATGTATGACAGCGCACTGGTGC	540
Db	528	GGTCTAGGAAGTCCAAATTAATGACTTTAGACCACAGACGATGTATGACAGCGCACTGGTGC	587
QY	541	CTTTTGGTCTTAAGAAGCCAGTATTAAGTGGCACAGTAAATGGTGAATACCTTCTCAAGC	600
Db	588	CTTTTGGTCTTAAGAAGCCAGTATTAAGTGGCACAGTAAATGGTGAATACCTTCTCAAGC	647
QY	601	AGAAAGTG--CAGTGAAGTGAACGA--GGTGGTTACAAAGGTTTAA----TGAAAG	652
Db	648	AGAAAGTGAGCAGTGAAGTGAACGACGTTGGTTCCCAAGGTTTACCTCGAAAGAAAG	707
QY	653	TAATTAACAGGCTCTGG--AAAGAAATCTTGG-----AAGTCAGAGCAGAAAGGAGAGA	704

Db	708	TAATTAACCAAGTCCGGGAAAAGAAATTTTGTGGGAAGTCCACAGAGAGAAAGAGGCACCA	767
Qy	705	AAGTAGTGATACTCAAGAGCAAAAGTGACC--TACATACCCCCTCTCCACCTGAGGA	761
Db	768	ACGATGATGATACCTCAAGGGGACCAAGTGACCTTAATTTACCCCTGCTTCACATGTAA	827
Qy	762	TGAGAGCTCATCTTTTGACATTAT	786
Db	828	TGAGAGACTCCACTTTTGGCAATTT	852

RESULT 15	LOCUS	DEFINITION
EG717812	657 bp	mrna
EG717812		linear
60269382EF1	NH_MGC_97	Homo sapiens cDNA clone IMAGE:4826120 5', 3'

RESULT	15
LOCUS	BG717812
DEFINITION	BG717812 657 bp mRNA linear EST 08-MAY-2001 602693829F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826190 5', mRNA sequence.
ACCESSION	BG717812
VERSION	BG717812.1 GI:13996999
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 657)
TITLE	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a>
JOURNAL	National Institutes of Health, "Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgabds-r@mail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MCC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.lnl.gov>  
Plate: L14M10740 row: 1 column: 07  
High quality sequence stop: 655.

```

FEATURES
Source
Location/Qualifiers
1. 657
/organism="Homo sapiens"
/db xref="taxon:9606"
/cclone="IMAGE:4826190"
/cclone_lib="NIH_MGC_97"
/lab_host="DH10E"
/notes="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI; XhoI (StuGag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTT-3',
size-selected for average insert size 2.2 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation) library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT
200 a 105 c 192 g 160 t

```

Query Match	27.3%	Score 593.4;	DB 12;	Length 657;
Best Local Similarity	99.7%	Pred. No. 3e-156;		
Matches 605;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;

QY 1 ATGGGGGATGAATGTTGGGAAGCAGAAATCAACCCGTATGTCCTCTATGTTCCCAT 60  
Db 51 ATGGGAGATACAGATTGGGAGCAGAAATACACCCGTATGTCCTCTATGTTCCCAT 110  
QY 61 TTTCAGAGAGATAGGTAATTCGTGAGAAAAATGAGACATTTTAAACAGACTCCAGCTTCA 120  
Db 111 TTTCAGAGAGATAGGTAATTCGTGAGAAAAATGAGACATTTTAAACAGACTCCAGCTTCA 170  
QY 121 TCATCAGAAATGGATATGAGCCTTCTCCGAAGATCATTTTCATGAAAAAGTGGATTTGGCC 180

```
Db      171 TCATCAGAAATGATGATGACCTTCTCGAAGAGATCATTTTCATGAAAAGTGATTTGCC 230
QY      181 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 240
Db      231 TCTGGCGGAATTTTGGAAACAGAGATGCTGTGTATTAAGCGAGATATACATCC 290
QY      241 ACAATGGGTGTTTGGAGTTGAAAGAGTTTGGAAAACAGAGTTTTCAAACAGCAGG 300
Db      291 ACAATGGGTGTTTGGAGTTGAAAGAGTTTGGAAAACAGAGTTTTCAAACAGCAGG 350
QY      301 TTGAAGATGATGATGATGCTGTGTTTCTGAGAGAGCTAGTAATGACTGGAAGATAAT 360
Db      351 TTGAAGATGATGATGATGCTGTGTTTCTGAGAGAGCTAGTAATGACTGGAAGATAAT 410
QY      361 CCAACAGGAAACAGAGGTTTCCAAAGAGGCGGCTATCGAGATGGAATTAATTCA-GA 419
Db      411 CCAACAGGAAACAGAGGTTTCCAAAGAGGCGGCTATCGAGATGGAATTAATTCA-GA 470
QY      420 AGCTTCAGGCGCATACAGAGAGGTGGAAGAGTAGTTCCGAGTTGCCGTGAGGATT 479
Db      471 AGCTTCAGGCGCATACAGAGAGGTGGAAGAGTAGTTCCGAGTTGCCGTGAGGATT 530
QY      480 TGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGATGTATGACGCGCACTGGTGG 539
Db      531 TGGTCTAGGAAGTCCAAATAATGACTTAGACCCAGACGATGTATGACGCGCACTGGTGG 590
QY      540 CCTTTTGGTCTAGAGACCAAGTATTAGTGCACAGGTATGTTGATTAATTCTCAAG 599
Db      591 CCTTTTGGTCTAGAGACCAAGTATTAGTGCACAGGTATGTTGATTAATTCTCTCAAG 650
QY      600 CAGAAAT 606
Db      651 CAGAAAT 657
```

Search completed: June 10, 2003, 10:40:19  
Job time : 2779 secs

IntelliGenetics

GENALIGN - Multiple Sequence Alignment Program  
Release 5.4

Wed 11 Jun 103 10:51:36-PDT

**Solution Parameters:**

**Solution Parameters:**

```
Nucleic Alphabet = Identity
Output line length = 80
Compress = Off
Histogram = Off
Randomization = Off
```

Output line length = 80

Compress	OF
History	OF

Randomization	Of
Randomized	Of

[illegible]

AMINO-Res-length = 2

Deletion-weight = 5.

Length-factor	= 0
Matching-weight	= 1

NUCLEIC-Res-length = 4

Spread-factor = 50

—

Clustered order of selected sequences:

1. US-09-714-865-1	(1-2224)
7. US-09-714-865-15	(1-2172)

US-09-714- 964 ATACTTGACGACGAGATTGATGGCTTGCTCAAAACAGGGCTGCGGAGACTCGGCTT  
consensus AATCTTGACGACGAGATTGATGGCTTGCTCAAAACAGGGCTGCGGAGACTCGGCTT  
US-09-714- 1038 TTCTCTCAACCAATTTTGGCTCATATGATGATGATAAATTAATCTCGATCGTTTAAAGA  
US-09-714- 1025 TTCTCTCAACCAATTTTGGCTCATATGATGATGATAAATTAATCTCGATCGTTTAAAGA  
consensus TTCTCTCAACCAATTTTGGCTCATATGATGATGATAAATTAATCTCGATCGTTTAAAGA  
US-09-714- 1099 GTTGACGAGAACGAGAGTATATTATGTAGACCACTCGAGATTGTCAACCAATTTAT  
US-09-714- 1086 GTTGACGAGAACGAGAGTATATTATGTAGACCACTCGAGATTGTCAACCAATTTAT  
consensus GTTGACGAGAACGAGAGTATATTATGTAGACCACTCGAGATTGTCAACCAATTTAT  
US-09-714- 1160 TTGGAAGCCAGAAATTTTCTTTGGACTTGTAAAGACTGTGTATATATATGAGGGA  
US-09-714- 1147 TTGGAAGCCAGAAATTTTCTTTGGACTTGTAAAGACTGTGTATATATATGAGGGA  
consensus TTGGAAGCCAGAAATTTTCTTTGGACTTGTAAAGACTGTGTATATATATGAGGGA  
US-09-714- 1221 CCCAGCTGGGACATTCATTCGACCAATAGTCAAGGCTGTAAATATATATGCTACTCC  
US-09-714- 1208 CCCAGCTGGGACATTCATTCGACCAATAGTCAAGGCTGTAAATATATATGCTACTCC  
consensus CCCAGCTGGGACATTCATTCGACCAATAGTCAAGGCTGTAAATATATATGCTACTCC  
US-09-714- 1282 TGGAGACTGATGATATATATAGGCAAGAAAGATTGGTCTCAACAGATCAAACTTA  
US-09-714- 1269 TGGAGACTGATGATATATATAGGCAAGAAAGATTGGTCTCAACAGATCAAACTTA  
consensus TGGAGACTGATGATATATATAGGCAAGAAAGATTGGTCTCAACAGATCAAACTTA  
US-09-714- 1343 GTTTTGGATGAAAGCTGATCGCATGTGTGATATGGTTTGGTCCAGAAATGAAGATTAA  
US-09-714- 1330 GTTTTGGATGAAAGCTGATCGCATGTGTGATATGGTTTGGTCCAGAAATGAAGATTAA  
consensus GTTTTGGATGAAAGCTGATCGCATGTGTGATATGGTTTGGTCCAGAAATGAAGATTAA  
US-09-714- 1404 TTTCTTGCCAGAGAAATGCCATCAAGAAACAGGCCAAACCTTATATGTTCAAGTCAACTT  
US-09-714- 1391 TTTCTTGCCAGAGAAATGCCATCAAGAAACAGGCCAAACCTTATATGTTCAAGTCAACTT  
consensus TTTCTTGCCAGAGAAATGCCATCAAGAAACAGGCCAAACCTTATATGTTCAAGTCAACTT  
US-09-714- 1465 TCCAGAGAAATTCAAAGGTGGCTGACAGATTTTAAAGTCAAAATTAATCTGTTGCT  
US-09-714- 1452 TCCAGAGAAATTCAAAGGTGGCTGACAGATTTTAAAGTCAAAATTAATCTGTTGCT  
consensus TCCAGAGAAATTCAAAGGTGGCTGACAGATTTTAAAGTCAAAATTAATCTGTTGCT  
US-09-714- 1526 GTTGACAGAGTGGGTGAGCATGTAGAGATGTTCAACAGACCGTTCTCAAGTTGGCCAGT  
US-09-714- 1513 GTTGACAGAGTGGGTGAGCATGTAGAGATGTTCAACAGACCGTTCTCAAGTTGGCCAGT  
consensus GTTGACAGAGTGGGTGAGCATGTAGAGATGTTCAACAGACCGTTCTCAAGTTGGCCAGT  
US-09-714- 1587 TCTCAAAAAGAGAAAGCTGTTGAATTTCTGGAAACATAGGGGATGAAGAACTATGCT  
US-09-714- 1574 TCTCAAAAAGAGAAAGCTGTTGAATTTCTGGAAACATAGGGGATGAAGAACTATGCT  
consensus TCTCAAAAAGAGAAAGCTGTTGAATTTCTGGAAACATAGGGGATGAAGAACTATGCT

US-09-714- 1648 CTTTGTGAACCTAAGAAAAAGCAGATTTTACTGCACTTTCTTGTCAAGAAAAATA  
US-09-714- 1635 CTTTGTGAACCTAAGAAAAAGCAGATTTTACTGCACTTTCTTGTCAAGAAAAATA  
consensus CTTTGTGAACCTAAGAAAAAGCAGATTTTACTGCACTTTCTTGTCAAGAAAAATA  
US-09-714- 1709 TCAACTACAGTATCCATGTGATCGGGAACAGAGAGCGGGAGCAAGCTCTTGAGATT  
US-09-714- 1696 TCAACTACAGTATCCATGTGATCGGGAACAGAGAGCGGGAGCAAGCTCTTGAGATT  
consensus TCAACTACAGTATCCATGTGATCGGGAACAGAGAGCGGGAGCAAGCTCTTGAGATT  
US-09-714- 1770 TTGCTTTGAAAGTGGCCAGTCTTGTGCTACTTCAATAGCTGCCAGAGGCTGATAT  
US-09-714- 1757 TTGCTTTGAAAGTGGCCAGTCTTGTGCTACTTCAATAGCTGCCAGAGGCTGATAT  
consensus TTGCTTTGAAAGTGGCCAGTCTTGTGCTACTTCAATAGCTGCCAGAGGCTGATAT  
US-09-714- 1831 TGAATATGTGCAACATGTTATCAATTTTGATCTTCTTCAACATGATGAATGTTTCA  
US-09-714- 1818 TGAATATGTGCAACATGTTATCAATTTTGATCTTCTTCAACATGATGAATGTTTCA  
consensus TGAATATGTGCAACATGTTATCAATTTTGATCTTCTTCAACATGATGAATGTTTCA  
US-09-714- 1892 CGAATTTGGGCGTACTGTCGTTGTGGGAATATCTGACAGCAATTTCTTTTGTATCTTG  
US-09-714- 1879 CGAATTTGGGCGTACTGTCGTTGTGGGAATATCTGACAGCAATTTCTTTTGTATCTTG  
consensus CGAATTTGGGCGTACTGTCGTTGTGGGAATATCTGACAGCAATTTCTTTTGTATCTTG  
US-09-714- 1953 AATCGATTAACATTTAGACAGGCTCTAGTAAAGATTGACAGATGCTCAACAGATGT  
US-09-714- 1940 AATCGATTAACATTTAGACAGGCTCTAGTAAAGATTGACAGATGCTCAACAGATGT  
consensus AATCGATTAACATTTAGACAGGCTCTAGTAAAGATTGACAGATGCTCAACAGATGT  
US-09-714- 2014 TCCGTCATGTTGGAAGAAATTTGCTTTAGTACATACATTCCTGGCTTCAAGTGTATACA  
US-09-714- 2001 TCCGTCATGTTGGAAGAAATTTGCTTTAGTACATACATTCCTGGCTTCAAGTGTATACA  
consensus TCCGTCATGTTGGAAGAAATTTGCTTTAGTACATACATTCCTGGCTTCAAGTGTATACA  
US-09-714- 2075 AGAGGAAACGTGTTTGCATCACTTGAATACAGAAAGGCAAGACCTTTGAACACAGCTG  
US-09-714- 2062 AGAGGAAACGTGTTTGCATCACTTGAATACAGAAAGGCAAGACCTTTGAACACAGCTG  
consensus AGAGGAAACGTGTTTGCATCACTTGAATACAGAAAGGCAAGACCTTTGAACACAGCTG  
US-09-714- 2136 GGTTTTCTTCTTCAACGAGCTCCCAATCCAGTATGATGATGATCATGGATCAAGCAAAA  
US-09-714- 2123 GGTTTTCTTCTTCAACGAGCTCCCAATCCAGTATGATGATGATCATGGATCAAGCAAAA  
consensus GGTTTTCTTCTTCAACGAGCTCCCAATCCAGTATGATGATGATCATGGATCAAGCAAAA  
US-09-714- 2197 catccttcaagtcggtgttgatgca  
US-09-714- 2173 catccttcaagtcggtgttgatgca  
consensus catccttcaagtcggtgttgatgca

Alignment score = 1912.00

Scoring matrix:

		1	7
1			
7		1912	

-----

-----

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

Page 1

ig. Frame

	0.00	
	0	
	0	
		70
		TCCCAT
		TCCCAT
		40
		AAGAAAT
		AAGAAAT
		130

AAAACAG  
 |||||  
 AAAACAG  
 200  
 TTTTGG  
 |||||  
 TTTTGG  
 0

360  
TTATGGA  
|||||  
TTATGGA

430  
ATTCAGA

0000

—  
—  
—  
—  
—  
—

```
ig. Frame
-----
0.00    0
```

0  
.  
0  
0 0 0

70	TCCCAT	40	AGAAAT
TCCCAT		AGAAAT	
		130	

AAAACAG  
 |||||  
 AAAACAG  
 200  
 TTTTGG  
 |||||  
 TTTTGG  
 0

360  
TTAATGA  
|||||  
TTAATGA

430  
ATTGAGA  
|||||  
ATTGAGA

100  
 TACGGAAG  
 |||||  
 TACGGAAG  
 490





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:54:21 ; Search time 74 Seconds  
(without alignments)  
12095.523 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 724  
Sequence: 1 atcggggagatgaagatcgga.....tagatgatgacatcgagat 2172

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Command line parameters:

-MODE=frame+ n2p.model -DRV=xlp  
-Q/cgn2\_1/USPTO.spcol/US09714865/runat\_05062003\_111833\_26583/app.query.fasta\_1.2311  
-DB=SEPREMBL\_21 -QFMT=fastean -SUFFIX=oligo.rspt -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=1000 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09714865.@CCN\_1\_1\_119@runat\_05062003\_111833\_26583 -NCPU=6 -ICPU=3  
-NO\_MMAP -LANG=QUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARYS

Result	Query	Score	Match	Length	DB	ID	Description
1	27	3.7	546	5	Q9GV10		Q9gv10 ephydactia f

2	25	3.5	394	13	Q8QGD0		Q8qgd0 pantodon bu
3	24	3.3	399	13	Q918L8		Q918l8 danio dangi
4	24	3.3	617	13	Q90ZP6		Q90zf6 oryzias lat
5	24	3.3	621	13	Q8QHL6		Q8qhl6 oreochromis
6	24	3.3	645	13	Q9DEB3		Q9deg3 oreochromis
7	24	3.3	647	13	Q9PT10		Q9pt10 oncorhynch
8	24	3.3	700	13	Q91372		Q91372 xenopus lae
9	24	3.3	700	13	Q42378		Q42378 brachydanio
10	24	3.3	715	13	Q8QCG8		Q8qcg8 brachydanio
11	24	3.3	715	13	Q8QFUD		Q8qfu0 brachydanio
12	24	3.3	716	13	Q42107		Q42107 brachydanio
13	22	3.0	400	11	Q8QGC9		Q8qgc9 melanotaeni
14	22	3.0	658	11	Q9QWS9		Q9qws9 mus musculu
15	21	2.9	396	13	Q8QGC8		Q8qgc8 oncorhynch
16	20	2.8	659	5	Q96069		Q96069 ciona intes
17	20	2.8	669	5	Q96068		Q96068 ciona intes
18	20	2.8	688	5	Q96067		Q96067 ciona intes
19	20	2.8	688	13	Q42375		Q42375 brachydanio
20	20	2.8	770	5	Q9GNP1		Q9gnp1 ciona savig
21	19	2.6	376	5	Q8SXU8		Q8sxu8 drosophila
22	19	2.6	396	13	Q8QGD1		Q8qgd1 oryzias lat
23	19	2.6	397	13	Q8QGD2		Q8qgd2 hyphessobry
24	19	2.6	398	13	Q8QGD3		Q8qgd3 cyprius ca
25	19	2.6	573	5	Q9GV07		Q9gv07 dugesia dor
26	19	2.6	646	10	Q9M2F9		Q9m2f9 arabidopsis
27	19	2.6	798	5	Q9VHP0		Q9vhp0 drosophila
28	19	2.6	798	5	Q8SX18		Q8sx18 drosophila
29	18	2.5	491	5	Q9GV11		Q9gv11 ephydactia f
30	18	2.5	628	5	Q9GV14		Q9gv14 hydra magni
31	18	2.5	633	10	Q9S1N6		Q9s1n6 arabidopsis
32	18	2.5	726	5	Q97031		Q97031 dugesia jap
33	18	2.5	763	5	Q22873		Q22873 caenorhabdi
34	18	2.5	974	5	Q96619		Q96619 caenorhabdi
35	18	2.5	974	5	Q27376		Q27376 caenorhabdi
36	17	2.3	603	10	Q9M2G5		Q9m2g5 arabidopsis
37	17	2.3	662	13	Q9PDG8		Q9pdg8 gallus galli
38	17	2.3	688	3	Q9PEU9		Q9peu9 neuropept
39	16	2.2	225	3	Q9UUI4		Q9ui4 schizosacch
40	16	2.2	617	3	Q8TFK8		Q8tfk8 candida gla
41	16	2.2	641	5	Q9N3F4		Q9n3f4 caenorhabdi
42	16	2.2	644	5	Q9SXY1		Q9sxy1 caenorhabdi
43	15	2.1	108	5	Q9NHW6		Q9nhw6 drosophila
44	15	2.1	255	3	Q93932		Q93932 candida gla
45	15	2.1	263	10	Q9ZRZ7		Q9zrz7 arabidopsis
46	15	2.1	328	5	Q9BL38		Q9bl38 caenorhabdi
47	15	2.1	480	10	Q8S0B0		Q8s0b0 oryza sativ
48	15	2.1	484	10	Q8W4R3		Q8w4r3 arabidopsis
49	15	2.1	494	10	Q941W4		Q941w4 oryza sativ
50	15	2.1	501	10	Q9C718		Q9c718 arabidopsis
51	15	2.1	561	5	Q9XUW5		Q9xuw5 caenorhabdi
52	15	2.1	564	10	Q9FMT1		Q9fml1 arabidopsis
53	15	2.1	578	5	Q917P5		Q917p5 drosophila
54	15	2.1	595	13	Q9W744		Q9w744 gallus galli
55	15	2.1	608	13	Q918S3		Q918s3 xenopus lae
56	15	2.1	618	5	Q9XTP2		Q9xtp2 molgula oca
57	15	2.1	618	10	Q93WJ3		Q93wj3 arabidopsis
58	15	2.1	619	5	Q9Y1M2		Q9y1m2 molgula occ
59	15	2.1	619	10	Q64430		Q64430 arabidopsis
60	15	2.1	619	10	Q8W4F5		Q8w4f5 arabidopsis
61	15	2.1	630	5	Q9NSK1		Q9nsk1 caenorhabdi
62	15	2.1	703	5	Q9VHV7		Q9vhv7 drosophila
63	15	2.1	713	10	Q9LYJ9		Q9lyj9 arabidopsis
64	15	2.1	716	10	Q9SP26		Q9sp26 pisum sativ
65	15	2.1	733	10	Q942R9		Q942r9 oryza sativ
66	15	2.1	781	5	Q97032		Q97032 dugesia jap
67	15	2.1	811	5	Q9VRS5		Q9vrz5 drosophila
68	15	2.1	891	5	Q8RYT6		Q8ryt6 oryza sativ
69	15	2.1	945	5	Q9W3M7		Q9w3m7 drosophila
70	14	1.9	421	2	Q93AA5		Q93aa5 versinia pe
71	14	1.9	421	10	Q8W4P3		Q8w4p3 arabidopsis
72	14	1.9	422	16	Q9KLB6		Q9klb6 vibrio chol
73	14	1.9	423	16	Q9KU63		Q9ku63 vibrio chol
74	14	1.9	425	2	Q9Z152		Q9z152 anabaena va

75	14	1.9	426	16	Q8YN52	Q8YN52 anabaena sp	148	11	1.5	517	16	Q9A779	Q9A779 caulobacter
76	14	1.9	446	16	Q9PCV6	Q9PCV6 xylella fas	149	11	1.5	527	16	Q9ENAT	Q9ENAT rhizobium s
77	14	1.9	446	16	Q9HYE0	Q9HYE0 pseudomonas	150	11	1.5	608	16	Q9EYU0	Q9EYU0 ralsstonia s
78	14	1.9	451	16	Q9KMW4	Q9KMW4 vibrio chol	151	11	1.5	610	16	Q9ELU3	Q9ELU3 pasteurella
79	14	1.9	454	16	Q8ZQW8	Q8ZQW8 salmonella	152	11	1.5	646	16	Q8ZL74	Q8ZL74 salmonella
80	14	1.9	455	16	Q8Z877	Q8Z877 salmonella	153	11	1.5	646	16	Q8Z311	Q8Z311 salmonella
81	14	1.9	455	16	Q8X7Y7	Q8X7Y7 escherichia	154	11	1.5	646	16	Q8XA87	Q8XA87 escherichia
82	14	1.9	458	16	Q916G3	Q916G3 pseudomonas	155	11	1.5	648	16	Q9NXZ2	Q9NXZ2 homo sapien
83	14	1.9	460	16	Q8ZFD8	Q8ZFD8 yersinia pe	156	11	1.5	662	3	Q9Y7T7	Q9Y7T7 schizosacch
84	14	1.9	462	16	Q9JY77	Q9JY77 neisseria m	157	11	1.5	664	16	Q8ZBC8	Q8ZBC8 yersinia pe
85	14	1.9	462	16	Q9JY73	Q9JY73 neisseria m	158	11	1.5	1016	5	Q9XTX5	Q9XTX5 caenorhabdi
86	14	1.9	465	16	Q8Y1R7	Q8Y1R7 ralsstonia s	159	11	1.5	1156	5	Q44758	Q44758 caenorhabdi
87	14	1.9	495	16	Q8SRB2	Q8SRB2 encephalito	160	10	1.4	110	5	Q01955	Q01955 bombyx mori
88	14	1.9	495	16	Q8XXA5	Q8XXA5 ralsstonia s	161	10	1.4	111	2	Q9L6R9	Q9L6R9 salmonella
89	14	1.9	522	16	Q9P9G7	Q9P9G7 methanococc	162	10	1.4	113	5	Q9GUT4	Q9GUT4 aurelia aur
90	14	1.9	535	16	Q8YH70	Q8YH70 bruceella me	163	10	1.4	113	5	Q9GUT7	Q9GUT7 sanderia ma
91	14	1.9	540	16	Q8Y1Z8	Q8Y1Z8 ralsstonia s	164	10	1.4	114	5	Q9GUT3	Q9GUT3 acioporta di
92	14	1.9	639	16	Q916B9	Q916B9 pseudomonas	165	10	1.4	114	5	Q9GUT2	Q9GUT2 craepedacus
93	14	1.9	732	5	Q9N478	Q9N478 caenorhabdi	166	10	1.4	114	5	Q9GUT0	Q9GUT0 hydractinia
94	14	1.9	1088	17	Q9SCV1	Q9SCV1 arabidopsis	167	10	1.4	114	5	Q9GUT9	Q9GUT9 tima formos
95	13	1.8	368	17	Q28030	Q28030 archaeoglob	168	10	1.4	115	5	Q01956	Q01956 bombyx mori
96	13	1.8	397	16	Q9KES3	Q9KES3 vibrio chol	169	10	1.4	123	5	Q01958	Q01958 bombyx mori
97	13	1.8	397	16	Q9HXE5	Q9HXE5 pseudomonas	170	10	1.4	155	10	Q9C8E6	Q9C8E6 arabidopsis
98	13	1.8	411	10	Q9ZS13	Q9ZS13 arabidopsis	171	10	1.4	187	10	Q9ZS02	Q9ZS02 arabidopsis
99	13	1.8	457	16	Q9JYU9	Q9JYU9 neisseria m	172	10	1.4	224	10	Q9ZS09	Q9ZS09 arabidopsis
100	13	1.8	457	16	Q9JYU9	Q9JYU9 neisseria m	173	10	1.4	240	4	Q9NT96	Q9NT96 homo sapien
101	13	1.8	463	5	Q9BL37	Q9BL37 caenorhabdi	174	10	1.4	376	10	Q9ZS10	Q9ZS10 arabidopsis
102	13	1.8	474	10	Q9SMW9	Q9SMW9 lycopersico	175	10	1.4	421	16	Q8XAT4	Q8XAT4 escherichia
103	13	1.8	484	5	Q9SKD7	Q9SKD7 caenorhabdi	176	10	1.4	423	16	Q9CJ51	Q9CJ51 pasteurella
104	13	1.8	504	5	Q9XVZ6	Q9XVZ6 caenorhabdi	177	10	1.4	425	17	Q26305	Q26305 melhanobact
105	13	1.8	537	10	Q9CS51	Q9CS51 arabidopsis	178	10	1.4	428	16	Q8ZAD8	Q8ZAD8 yersinia pe
106	13	1.8	601	5	Q01378	Q01378 bombyx mori	179	10	1.4	438	16	Q9KVS2	Q9KVS2 vibrio chol
107	13	1.8	622	10	Q9LKL6	Q9LKL6 pium sativ	180	10	1.4	456	10	Q9FMEL	Q9FMEL arabidopsis
108	13	1.8	747	5	Q96SK1	Q96SK1 caenorhabdi	181	10	1.4	460	16	Q9KPI5	Q9KPI5 vibrio chol
109	13	1.8	797	5	Q9GVJ3	Q9GVJ3 hydra magni	182	10	1.4	470	16	Q9ZCF8	Q9ZCF8 listeria in
110	13	1.8	970	5	Q96SK2	Q96SK2 caenorhabdi	183	10	1.4	470	16	Q8Y7M8	Q8Y7M8 listeria mo
111	13	1.7	202	2	Q93873	Q93873 xanthomonas	184	10	1.4	478	16	Q8XW28	Q8XW28 clostridium
112	13	1.7	206	2	P74957	P74957 shewanella	185	10	1.4	497	2	Q9Z6C9	Q9Z6C9 lactobacilli
113	13	1.7	337	11	Q9CSD1	Q9CSD1 mus muscucu	186	10	1.4	502	16	Q8UCZ5	Q8UCZ5 agrobacteri
114	13	1.7	366	5	Q9U936	Q9U936 plasmidium	187	10	1.4	503	16	Q9ZUF7	Q9ZUF7 rhizobium m
115	13	1.7	411	2	Q87951	Q87951 vibrio algi	188	10	1.4	511	16	P96614	P96614 bacillus su
116	13	1.7	441	16	Q8ZD82	Q8ZD82 yersinia pe	189	10	1.4	513	16	Q9KXJ0	Q9KXJ0 anabaena sp
117	13	1.7	444	16	Q8ZMX7	Q8ZMX7 salmonella	190	10	1.4	516	16	Q9ZDF4	Q9ZDF4 listeria in
118	13	1.7	444	16	Q8Z4J8	Q8Z4J8 salmonella	191	10	1.4	520	16	Q8XW80	Q8XW80 listeria mo
119	13	1.7	444	16	Q8XAX21	Q8XAX21 escherichia	192	10	1.4	522	2	Q93HQ5	Q93HQ5 streptococc
120	13	1.7	446	2	Q8RT7	Q8RT7 uncultured	193	10	1.4	527	16	Q05855	Q05855 mycobacteri
121	13	1.7	542	10	Q9SZB4	Q9SZB4 arabidopsis	194	10	1.4	528	16	Q97EB6	Q97EB6 clostridium
122	13	1.7	580	4	Q9H0W3	Q9H0W3 homo sapien	195	10	1.4	528	16	Q8XKJ8	Q8XKJ8 clostridium
123	13	1.7	591	10	Q9L046	Q9L046 arabidopsis	196	10	1.4	539	16	Q8XHN6	Q8XHN6 fiobacteri
124	13	1.7	619	11	Q9DBN9	Q9DBN9 mus muscucu	197	10	1.4	539	16	Q9KAA6	Q9KAA6 bacillus ha
125	13	1.7	626	11	Q8R0R6	Q8R0R6 mus muscucu	198	10	1.4	539	16	Q9ZJ38	Q9ZJ38 streptococc
126	13	1.7	730	5	Q95QW2	Q95QW2 caenorhabdi	199	10	1.4	544	16	Q9CCH3	Q9CCH3 mycobacteri
127	13	1.7	769	5	Q95QW3	Q95QW3 caenorhabdi	200	10	1.4	547	16	Q9CIR6	Q9CIR6 lactococcus
128	13	1.7	844	5	Q9NNF5	Q9NNF5 leishmania	201	10	1.4	555	17	Q8TJ32	Q8TJ32 methanosarc
129	13	1.7	890	5	Q9GV12	Q9GV12 hydra magni	202	10	1.4	567	16	Q91003	Q91003 pseudomonas
130	13	1.7	1031	4	Q9Y658	Q9Y658 homo sapien	203	10	1.4	578	3	Q9Y7C4	Q9Y7C4 candida alb
131	13	1.7	1031	4	Q96ER0	Q96ER0 homo sapien	204	10	1.4	589	10	Q91GMR4	Q91GMR4 oryza sativ
132	13	1.7	1032	4	Q94894	Q94894 homo sapien	205	10	1.4	610	10	Q9FZ92	Q9FZ92 arabidopsis
133	13	1.7	1032	11	Q6Z780	Q6Z780 rattus norv	206	10	1.4	614	16	Q9PFP6	Q9PFP6 xylella fas
134	13	1.5	121	5	Q9GV01	Q9GV01 aurelia aur	207	10	1.4	622	11	Q91VW6	Q91VW6 mus muscucu
135	13	1.5	122	5	Q9GV06	Q9GV06 sanderia ma	208	10	1.4	680	5	Q95TJ3	Q95TJ3 drosophila
136	13	1.5	343	17	Q97X64	Q97X64 sulfolobus	209	10	1.4	760	10	Q22907	Q22907 arabidopsis
137	13	1.5	360	17	Q97WTO	Q97WTO sulfolobus	210	10	1.4	763	5	Q95XW9	Q95XW9 caenorhabdi
138	13	1.5	449	16	Q9HX66	Q9HX66 pseudomonas	211	10	1.4	820	4	Q43188	Q43188 homo sapien
139	13	1.5	452	16	Q9KVA4	Q9KVA4 vibrio chol	212	10	1.4	820	4	Q9BU08	Q9BU08 homo sapien
140	13	1.5	457	16	Q8ZP87	Q8ZP87 salmonella	213	10	1.4	845	10	Q49289	Q49289 arabidopsis
141	13	1.5	457	16	Q8X8Q4	Q8X8Q4 escherichia	214	10	1.4	855	5	P90549	P90549 leishmania
142	13	1.5	465	3	Q9P6M8	Q9P6M8 schizosacch	215	10	1.4	989	10	Q9SFF41	Q9SFF41 arabidopsis
143	13	1.5	467	16	Q98N14	Q98N14 rhizobium l	216	10	1.4	1014	3	Q9P7C7	Q9P7C7 schizosacch
144	13	1.5	482	16	Q8YCX0	Q8YCX0 bruceella me	217	10	1.4	1166	10	Q9SYP6	Q9SYP6 arabidopsis
145	13	1.5	491	10	Q9SAZ7	Q9SAZ7 arabidopsis	218	10	1.4	1224	5	Q9YXW2	Q9YXW2 drosophila
146	13	1.5	499	16	Q9ZPK3	Q9ZPK3 rhizobium m	219	9	1.2	49	2	Q9EYX4	Q9EYX4 salmonella
147	11	1.5	503	16	Q8UEC7	Q8UEC7 agrobacteri	220	9	1.2	93	3	Q9USB9	Q9USB9 schizosacch

221	9	1.2	106	5	097029	097029 dugesia jap	C 294	9	1.3	902	5	085X13	085X13 drosophila
222	9	1.2	107	5	097026	097026 dugesia jap	C 295	9	1.3	1152	5	095Y5	095Y5 leishmania
223	9	1.2	108	5	097021	097021 dugesia jap	C 296	9	1.3	1884	5	09NHM2	09NHM2 nephila mad
224	9	1.2	108	5	097030	097030 dugesia jap	C 297	9	1.3	2205	10	08RYW8	08RYW8 oryza sativ
225	9	1.2	110	5	097022	097022 dugesia jap	C 298	9	1.3	2336	5	09NGO2	09NGO2 dictyosteli
226	9	1.3	111	13	09YH66	09YH66 brachydanio	C 299	9	1.3	3112	5	09NKP1	09NKP1 leishmania
227	9	1.2	114	5	002028	002028 bombyx mori	C 300	9	1.3	3469	5	09U412	09U412 drosophila
228	9	1.3	162	10	09MOB4	09MOB4 arbidops	C 301	9	1.3	3502	5	09YVJ9	09YVJ9 drosophila
229	9	1.2	166	4	09EVQ8	09EVQ8 homo sapien	C 302	9	1.1	78	12	09YVY2	09YVY2 melanococ
230	9	1.2	166	4	09EVK4	09EVK4 rhizobium 1	C 303	8	1.1	71	2	054854	054854 streptococ
231	9	1.3	228	10	081323	081323 arbidops	C 304	8	1.1	92	4	09HVZ7	09HVZ7 homo sapien
232	9	1.2	257	4	09NVY6	09NVY6 homo sapien	C 305	8	1.1	100	17	097MW9	097MW9 sulfolobus
233	9	1.3	262	5	09VIX5	09VIX5 drosophila	C 306	8	1.1	110	5	097025	097025 dugesia jap
234	9	1.3	287	5	09VIX6	09VIX6 drosophila	C 307	8	1.1	110	5	097027	097027 dugesia jap
235	9	1.3	301	10	08S2F7	08S2F7 oryza sativ	C 308	8	1.1	111	6	077725	077725 gorilla gor
236	9	1.2	312	4	08WYGA	08WYGA homo sapien	C 309	8	1.1	111	6	077731	077731 pan troglod
237	9	1.3	323	12	067714	067714 human adeno	C 310	8	1.1	111	6	077703	077703 canis famil
238	9	1.3	323	12	067733	067733 human adeno	C 311	8	1.1	111	11	088757	088757 ractus norv
239	9	1.3	323	12	091CL7	091CL7 human adeno	C 312	8	1.1	113	5	001957	001957 bombyx mori
240	9	1.2	375	4	096F57	096F57 homo sapien	C 313	8	1.1	113	5	09GV08	09GV08 dugesia dor
241	9	1.2	375	17	08T163	08T163 methanosarc	C 314	8	1.1	114	5	09GUZ1	09GUZ1 etrene sp.
242	9	1.2	376	16	034750	034750 bacillus su	C 315	8	1.1	114	10	09MAW9	09MAW9 oryza sativ
243	9	1.3	378	5	076736	076736 dictyosteli	C 316	8	1.1	114	10	09MAW8	09MAW8 oryza sativ
244	9	1.2	406	4	09NMU3	09NMU3 homo sapien	C 317	8	1.1	117	5	09GV09	09GV09 dugesia dor
245	9	1.3	406	5	09NG14	09NG14 brachlosto	C 318	8	1.1	118	5	09GV00	09GV00 actopora di
246	9	1.3	412	3	09P603	09P603 neurospora	C 319	8	1.1	119	5	09GUZ7	09GUZ7 hydractinia
247	9	1.2	441	16	08V6G5	08V6G5 listeria mo	C 320	8	1.1	120	5	09GUZ9	09GUZ9 craspedacus
248	9	1.2	442	16	09CUZ8	09CUZ8 pasteurrella	C 321	8	1.1	120	5	09GUZ8	09GUZ8 etrene sp.
249	9	1.2	442	16	092AT6	092AT6 listeria in	C 322	8	1.1	120	5	09GUZ5	09GUZ5 tima formos
250	9	1.3	448	5	09VYK1	09VYK1 drosophila	C 323	8	1.1	122	16	08YV51	08YV51 anabaena sp
251	9	1.2	453	3	09HGMS	09HGMS schizosacch	C 324	8	1.1	126	4	016273	016273 homo sapien
252	9	1.3	453	11	064167	064167 mus musculu	C 325	8	1.1	129	10	082295	082295 arbidops
253	9	1.2	472	5	096967	096967 drosophila	C 326	8	1.1	131	11	09DA59	09DA59 white spot
254	9	1.2	476	13	098TR4	098TR4 fuqu rubrip	C 327	8	1.1	142	2	093S56	093S56 mus musculu
255	9	1.2	481	16	097F19	097F19 clostridium	C 328	8	1.1	145	5	09Y722	09Y722 drosophila
256	9	1.2	483	4	09Y6V7	09Y6V7 homo sapien	C 329	8	1.1	147	12	08VAS8	08VAS8 white spot
257	9	1.2	486	5	09VKE2	09VKE2 drosophila	C 330	8	1.1	159	5	09NAGO	09NAGO caenorhabdi
258	9	1.2	503	5	08S6G7	08S6G7 encephalito	C 331	8	1.1	169	11	035662	035662 mus musculu
259	9	1.3	508	11	093UJ6	093UJ6 mus musculu	C 332	8	1.1	174	5	001544	001544 caenorhabdi
260	9	1.2	517	3	08X1T4	08X1T4 emericeila	C 333	8	1.1	176	16	09RJ74	09RJ74 streptomyc
261	9	1.3	543	16	08XW22	08XW22 raietonia s	C 334	8	1.1	177	4	09EBW8	09EBW8 homo sapien
262	9	1.3	548	5	09W3D0	09W3D0 drosophila	C 335	8	1.1	178	16	09RU05	09RU05 streptomyc
263	9	1.3	556	10	09LGY7	09LGY7 oryza sativ	C 336	8	1.1	188	4	060377	060377 homo sapien
264	9	1.3	558	13	09PUB5	09PUB5 brachydanio	C 337	8	1.1	189	10	09FO88	09FO88 nicotiana t
265	9	1.2	561	4	09BXFO	09BXFO homo sapien	C 338	8	1.1	191	10	09FO83	09FO83 nicotiana t
266	9	1.2	572	11	0921N6	0921N6 mus musculu	C 339	8	1.1	194	16	08RFY9	08RFY9 fusobacteri
267	9	1.2	582	5	023509	023509 dictyosteli	C 340	8	1.1	195	11	09WU76	09WU76 mus musculu
268	9	1.3	606	16	08REH6	08REH6 fusobacteri	C 341	8	1.1	214	5	09N7D9	09N7D9 leishmania
269	9	1.3	609	10	09ARZ1	09ARZ1 oryza sativ	C 342	8	1.1	224	10	08W0N9	08W0N9 oryza sativ
270	9	1.2	625	5	09NE37	09NE37 dictyosteli	C 343	8	1.1	226	4	09H4A8	09H4A8 homo sapien
271	9	1.3	684	5	08TR48	08TR48 dictyosteli	C 344	8	1.1	226	10	09FO87	09FO87 nicotiana t
272	9	1.3	691	5	09NKR7	09NKR7 leishmania	C 345	8	1.1	229	15	09DKV7	09DKV7 caprine art
273	9	1.2	696	5	09NSM1	09NSM1 caenorhabdi	C 346	8	1.1	242	10	094DU3	094DU3 oryza sativ
274	9	1.2	709	4	075619	075619 homo sapien	C 347	8	1.1	251	10	08W2W7	08W2W7 oryza sativ
275	9	1.2	733	10	093G08	093G08 arbidops	C 348	8	1.1	259	10	092S05	092S05 arbidops
276	9	1.2	736	10	094GY4	094GY4 oryza sativ	C 349	8	1.1	263	8	09B8G8	09B8G8 heterodopus
277	9	1.2	746	5	091340	091340 caenorhabdi	C 350	8	1.1	264	10	09FSL0	09FSL0 oryza sativ
278	9	1.2	758	10	08SAX7	08SAX7 oryza sativ	C 351	8	1.1	267	13	057652	057652 fuqu rubrip
279	9	1.2	765	4	096SN7	096SN7 homo sapien	C 352	8	1.1	267	16	08Z073	08Z073 anabaena sp
280	9	1.2	765	4	096GQ7	096GQ7 homo sapien	C 353	8	1.1	268	10	09SBR6	09SBR6 arbidops
281	9	1.3	781	11	089090	089090 mus musculu	C 354	8	1.1	273	11	0920V5	0920V5 ractus norv
282	9	1.2	782	5	09VJ78	09VJ78 drosophila	C 355	8	1.1	274	6	09BG12	09BG12 bos taurus
283	9	1.3	784	11	089087	089087 mus musculu	C 356	8	1.1	274	16	095019	095019 mycobacteri
284	9	1.2	789	10	092RZ8	092RZ8 arbidops	C 357	8	1.1	280	11	091VQ2	091VQ2 mus musculu
285	9	1.2	791	5	09VTC1	09VTC1 drosophila	C 358	8	1.1	283	5	09NB39	09NB39 plasmodium
286	9	1.2	796	4	096LU7	096LU7 homo sapien	C 359	8	1.1	289	10	09FO85	09FO85 nicotiana t
287	9	1.3	809	10	043715	043715 plium sativ	C 360	8	1.1	290	10	065659	065659 arbidops
288	9	1.2	811	5	093382	093382 caenorhabdi	C 361	8	1.1	294	10	09FO89	09FO89 nicotiana t
289	9	1.3	812	10	088RT4	088RT4 oryza sativ	C 362	8	1.1	296	11	09AUM8	09AUM8 oryza sativ
290	9	1.2	819	4	096BK1	096BK1 homo sapien	C 363	8	1.1	296	11	099X76	099X76 mus musculu
291	9	1.2	853	5	09N2X0	09N2X0 caenorhabdi	C 364	8	1.1	296	11	09CXH8	09CXH8 mus musculu
292	9	1.3	871	5	044358	044358 nephila cia	C 365	8	1.1	296	11	064012	064012 mus musculu
293	9	1.2	899	5	095XS9	095XS9 caenorhabdi	C 366	8	1.1	299	2	09AMJ7	09AMJ7 pseudomonas

C 367	8	1.1	299	2	Q9AM06	Q9am16 pseudomonas	C 440	8	1.1	457	10	Q949W1	Q949w1 arabidopsis
C 368	8	1.1	299	2	Q87489	Q87489 pseudomonas	C 441	8	1.1	461	10	Q9LTW4	Q9ltw4 arabidopsis
C 369	8	1.1	304	10	Q82321	Q82321 arabidopsis	C 442	8	1.1	461	16	Q9CK50	Q9ck50 pasteurella
C 370	8	1.1	304	11	Q9E0N6	Q9egn6 mus musculi	C 443	8	1.1	464	8	Q9B8C8	Q9b8c8 candida alb
C 371	8	1.1	305	8	Q02677	Q02677 podospira a	C 444	8	1.1	464	10	Q9VTS6	Q9vts6 arabidopsis
C 372	8	1.1	307	5	Q9VAC9	Q9vac9 drosophila	C 445	8	1.1	465	5	Q8SRN8	Q8srn8 encephalito
C 373	8	1.1	320	10	Q23160	Q23160 arabidopsis	C 446	8	1.1	465	10	Q9FVU4	Q9fvu4 arabidopsis
C 374	8	1.1	323	12	Q67712	Q67712 human adeno	C 447	8	1.1	466	10	Q9FE48	Q9fe48 nicotiana t
C 375	8	1.1	323	16	Q99UK0	Q99uk0 staphylococ	C 448	8	1.1	466	10	Q9FO92	Q9fo92 arabidopsis
C 376	8	1.1	323	16	Q8YTP3	Q8ytp3 anabaena sp	C 449	8	1.1	467	4	Q9BWT8	Q9bwt8 nicotiana t
C 377	8	1.1	324	10	Q9F084	Q9fgr4 nicotiana t	C 450	8	1.1	471	4	Q9HCH7	Q9hch7 homo sapien
C 378	8	1.1	324	10	Q942K9	Q942k9 oryza sativ	C 451	8	1.1	472	5	Q9SPU7	Q9spu7 chromomus
C 379	8	1.1	327	10	Q94D12	Q94d12 oryza sativ	C 452	8	1.1	473	5	Q9Y0S9	Q9yos9 drosophila
C 380	8	1.1	328	2	Q9L6R8	Q9l6r8 salmonella	C 453	8	1.1	474	10	Q22679	Q22679 arabidopsis
C 381	8	1.1	335	2	Q8VUZ9	Q8vuz9 streptomyce	C 454	8	1.1	476	16	Q9A9X6	Q9a9x6 caulobacter
C 382	8	1.1	337	17	Q96XQ7	Q96xq7 sulfolobus	C 455	8	1.1	477	10	Q65255	Q65255 arabidopsis
C 383	8	1.1	341	16	Q9RTY3	Q9rtc3 deinococcus	C 456	8	1.1	478	4	Q9HBZ9	Q9hbz9 homo sapien
C 384	8	1.1	345	16	Q92WS8	Q92ws8 rhizobium m	C 457	8	1.1	478	4	Q9NNU7	Q9nnu7 homo sapien
C 385	8	1.1	345	16	Q8YB95	Q8yb95 brucella me	C 458	8	1.1	478	11	Q921R0	Q921r0 mus musculi
C 386	8	1.1	346	16	Q8U6Y0	Q8u6y0 agrobacteri	C 459	8	1.1	479	11	Q9R3C7	Q9r3c7 mus musculi
C 387	8	1.1	346	16	Q9ACU5	Q9acu5 streptomyce	C 460	8	1.1	480	16	Q9BR80	Q9br80 mycoplasma
C 388	8	1.1	351	16	Q9EP30	Q9ep30 xyloella fas	C 461	8	1.1	483	3	Q8X0X2	Q8x0x2 neurospora
C 389	8	1.1	353	4	Q8TEC9	Q8tec9 homo sapien	C 462	8	1.1	483	5	Q8T0P0	Q8t0p0 drosophila
C 390	8	1.1	362	10	Q9ZS03	Q9zso3 arabidopsis	C 463	8	1.1	483	13	Q9DGP9	Q9dgp9 xenopus lae
C 391	8	1.1	362	10	Q9EJF8	Q9efj8 arabidopsis	C 464	8	1.1	487	2	Q9F0Z5	Q9f0z5 synchococc
C 392	8	1.1	365	4	Q9Y3V8	Q9y3v8 homo sapien	C 465	8	1.1	488	5	Q9NBA7	Q9nba7 plasmodium
C 393	8	1.1	365	16	Q9FQ86	Q9fq86 nicotiana t	C 466	8	1.1	491	4	Q9NVE0	Q9nve0 homo sapien
C 394	8	1.1	368	16	Q92P72	Q92p72 rhizobium m	C 467	8	1.1	491	16	Q9ZMT8	Q9zmt8 helicobacte
C 395	8	1.1	369	4	Q9UHL0	Q9uhl0 homo sapien	C 468	8	1.1	492	5	Q9NMA4	Q9nma4 plasmodium
C 396	8	1.1	369	11	Q9QY16	Q9qy16 rattus norv	C 469	8	1.1	492	16	Q25029	Q25029 helicobacte
C 397	8	1.1	369	11	Q9QY15	Q9qy15 mus musculi	C 470	8	1.1	494	10	Q55804	Q55804 synchocyst
C 398	8	1.1	369	11	Q8R1B6	Q8rib6 mus musculi	C 471	8	1.1	494	16	Q58RW8	Q58rw8 arabidopsis
C 399	8	1.1	370	4	Q96KE7	Q96ke7 homo sapien	C 472	8	1.1	496	10	Q9S289	Q9s289 arabidopsis
C 400	8	1.1	372	11	Q9D7Z2	Q9d7z2 mus musculi	C 473	8	1.1	496	16	Q9ARC0	Q9arc0 deinococcus
C 401	8	1.1	372	11	Q8R0U2	Q8rou2 mus musculi	C 474	8	1.1	498	10	Q9ZS08	Q9zso8 arabidopsis
C 402	8	1.1	372	13	Q90Y50	Q90y50 brachydanio	C 475	8	1.1	498	10	Q9M2E0	Q9m2e0 arabidopsis
C 403	8	1.1	373	17	Q978T9	Q978t9 thermoplaem	C 476	8	1.1	499	10	Q65275	Q65275 arabidopsis
C 404	8	1.1	374	16	Q97KY2	Q97ky2 clostridium	C 477	8	1.1	499	11	Q9CS87	Q9cs87 mus musculi
C 405	8	1.1	375	5	Q9NGF5	Q9ngf5 drosophila	C 478	8	1.1	502	12	Q8UYJ5	Q8uyj5 camelopox vi
C 406	8	1.1	375	5	Q9NGP4	Q9ngf4 drosophila	C 479	8	1.1	505	10	Q9ZS12	Q9zsl2 arabidopsis
C 407	8	1.1	375	5	Q9NGF3	Q9ngf3 drosophila	C 480	8	1.1	505	10	Q8RXK6	Q8rxk6 arabidopsis
C 408	8	1.1	375	5	Q9N6A0	Q9n6a0 drosophila	C 481	8	1.1	506	16	Q9SSH6	Q9ssh6 staphylococ
C 409	8	1.1	378	8	Q94ZK8	Q94zk8 cebus albit	C 482	8	1.1	507	10	Q94ZN1	Q94zn1 oryza sativ
C 410	8	1.1	379	5	Q44879	Q44879 caenorhabdi	C 483	8	1.1	510	2	Q07897	Q07897 thermus the
C 411	8	1.1	381	8	Q9SEV5	Q9sev5 guillardia	C 484	8	1.1	513	13	Q9DEG2	Q9deg2 gallus gall
C 412	8	1.1	384	5	Q8SYB8	Q8syb8 drosophila	C 485	8	1.1	520	5	Q9BLY0	Q9bly0 ciona intes
C 413	8	1.1	388	17	Q9HLD1	Q9hld1 thermoplaem	C 486	8	1.1	521	10	Q9FVY5	Q9fvy5 oryza sativ
C 414	8	1.1	389	10	Q94HS3	Q94hs3 oryza sativ	C 487	8	1.1	522	10	Q8YFG8	Q8yfg8 arabidopsis
C 415	8	1.1	390	10	Q9F090	Q9fg90 nicotiana t	C 488	8	1.1	524	5	Q9VFO3	Q9vfo3 drosophila
C 416	8	1.1	394	17	Q8TQ10	Q8tq10 methanosarc	C 489	8	1.1	526	4	Q8TBR3	Q8tbr3 homo sapien
C 417	8	1.1	398	11	Q9D8Q2	Q9d8q2 mus musculi	C 490	8	1.1	528	10	Q8O838	Q8o838 arabidopsis
C 418	8	1.1	399	4	Q9NTR9	Q9ntr9 homo sapien	C 491	8	1.1	528	10	Q94BVA	Q94bv4 arabidopsis
C 419	8	1.1	400	5	Q24731	Q24731 drosophila	C 492	8	1.1	528	11	Q9CS10	Q9cs10 mus musculi
C 420	8	1.1	407	16	Q9ZCQ0	Q9zcq0 rickettsia	C 493	8	1.1	528	11	Q8R4H7	Q8r4h7 mus musculi
C 421	8	1.1	414	16	Q9Z6V2	Q9z6v2 rickettsia	C 494	8	1.1	535	3	Q13622	Q13622 schizosacch
C 422	8	1.1	415	10	Q9FXZ6	Q9fxz6 arabidopsis	C 495	8	1.1	543	4	Q9PPI9	Q9ppi9 homo sapien
C 423	8	1.1	419	13	Q8OC81	Q8oc81 xenopus lae	C 496	8	1.1	544	5	Q61815	Q61815 caenorhabdi
C 424	8	1.1	420	4	Q9Y482	Q9y482 homo sapien	C 497	8	1.1	544	16	Q9PA24	Q9pa24 xyloella fas
C 425	8	1.1	420	13	Q9DEG1	Q9deg1 gallus gall	C 498	8	1.1	546	5	Q61097	Q61097 trypanosoma
C 426	8	1.1	423	10	Q9CS31	Q9cs31 arabidopsis	C 499	8	1.1	546	5	Q15919	Q15919 trypanosoma
C 427	8	1.1	425	10	Q9FQ91	Q9fq91 nicotiana t	C 500	8	1.1	550	13	Q93263	Q93263 gallus gall
C 428	8	1.1	426	4	Q8WV76	Q8wv76 homo sapien	C 501	8	1.1	552	5	Q9VNS2	Q9vns2 drosophila
C 429	8	1.1	428	16	Q9KKW0	Q9kkw0 vibrio chol	C 502	8	1.1	552	10	Q94LG7	Q94lg7 oryza sativ
C 430	8	1.1	438	10	Q9CAU2	Q9cau2 arabidopsis	C 503	8	1.1	557	3	Q93981	Q93981 gibberella
C 431	8	1.1	439	10	Q8VYMS	Q8vym5 arabidopsis	C 504	8	1.1	558	10	Q9C9H4	Q9c9h4 arabidopsis
C 432	8	1.1	441	10	Q9CAU4	Q9cau4 arabidopsis	C 505	8	1.1	561	11	Q9CXH6	Q9cxh6 mus musculi
C 433	8	1.1	441	10	Q9CAU3	Q9cau3 arabidopsis	C 506	8	1.1	566	5	Q23910	Q23910 dictyosteli
C 434	8	1.1	442	4	Q9HA19	Q9ha19 homo sapien	C 507	8	1.1	568	10	Q9LIH9	Q9lih9 arabidopsis
C 435	8	1.1	448	4	Q9H0U0	Q9h0u0 homo sapien	C 508	8	1.1	568	11	Q9D2M2	Q9d2m2 mus musculi
C 436	8	1.1	450	10	Q23Z51	Q23z51 arabidopsis	C 509	8	1.1	574	3	Q93990	Q93990 candida alb
C 437	8	1.1	451	10	Q9ZS07	Q9zso7 arabidopsis	C 510	8	1.1	588	10	Q9LT69	Q9lt69 arabidopsis
C 438	8	1.1	453	16	P72834	P72834 synchocyst	C 511	8	1.1	591	10	Q94ZN2	Q94zn2 oryza sativ
C 439	8	1.1	457	10	Q9FP87	Q9ff87 arabidopsis	C 512	8	1.1	593	10	Q9FLB0	Q9flb0 arabidopsis

513	8	1.1	594	5	Q9VVK8	Q9VVK8 drosophila	c 586	8	1.1	948	5	Q8T065	Q8T065 drosophila
514	8	1.1	595	3	Q60080	Q60080 schizosacch	587	8	1.1	953	4	Q14151	Q14151 homo sapien
515	8	1.1	598	4	Q9Y2R4	Q9Y2R4 homo sapien	c 588	8	1.1	965	10	Q9MA56	Q9MA56 arabidopsis
516	8	1.1	598	11	Q99PT0	Q99PT0 rattus norv	589	8	1.1	1005	5	Q95ZC5	Q95ZC5 leishmania
517	8	1.1	600	3	Q93935	Q93935 paracoccidi	590	8	1.1	1022	5	Q22308	Q22308 caenorhabd
518	8	1.1	603	5	Q9VM86	Q9VM86 drosophila	c 591	8	1.1	1096	5	Q9W117	Q9W117 drosophila
519	8	1.1	604	3	Q74491	Q74491 schizosacch	c 592	8	1.1	1108	5	Q9V788	Q9V788 drosophila
520	8	1.1	606	3	Q74764	Q74764 schizosacch	c 593	8	1.1	1164	5	Q9V161	Q9V161 drosophila
521	8	1.1	613	5	Q62372	Q62372 caenorhabd	c 594	8	1.1	1272	5	Q95SG4	Q95SG4 drosophila
522	8	1.1	613	5	Q9VHU1	Q9VHU1 drosophila	c 595	8	1.1	1283	5	Q8T5H0	Q8T5H0 anopheles g
523	8	1.1	615	16	Q8UBC1	Q8UBC1 agrobacteri	c 596	8	1.1	1283	5	Q8T5H0	Q8T5H0 anopheles g
524	8	1.1	618	13	Q90W76	Q90W76 oncorhynch	c 597	8	1.1	1338	5	Q9V416	Q9V416 drosophila
525	8	1.1	620	16	Q98E62	Q98E62 rhizobium l	c 598	8	1.1	1368	5	Q9V4H9	Q9V4H9 drosophila
526	8	1.1	621	10	Q941V5	Q941V5 oryza sativ	c 599	8	1.1	1432	10	Q9FPR8	Q9FPR8 chlamydom
527	8	1.1	626	10	Q9SM44	Q9SM44 arabidopsis	c 600	8	1.1	1443	5	Q9N8S1	Q9N8S1 trypanosoma
528	8	1.1	626	16	Q8YEQ4	Q8YEQ4 bruceella me	c 601	8	1.1	1458	5	Q8T9F7	Q8T9F7 drosophila
529	8	1.1	630	16	Q92SG6	Q92SG6 rhizobium m	602	8	1.1	1458	11	Q91W70	Q91W70 mus musculu
530	8	1.1	638	5	Q18056	Q18056 caenorhabd	603	8	1.1	1458	11	Q8VH06	Q8VH06 mus musculu
531	8	1.1	641	5	Q9S5F2	Q9S5F2 drosophila	c 604	8	1.1	1478	17	Q8TLD5	Q8TLD5 mechanosarc
532	8	1.1	643	5	Q8STR2	Q8STR2 dicystosteli	605	8	1.1	1484	4	Q96QH3	Q96QH3 homo sapien
533	8	1.1	647	10	Q94Z00	Q94Z00 oryza sativ	c 607	8	1.1	1487	11	Q9J1A1	Q9J1A1 mus musculu
534	8	1.1	648	3	Q74393	Q74393 schizosacch	c 608	8	1.1	1561	6	Q9GKPI	Q9GKPI sus scrofa
535	8	1.1	659	5	Q94600	Q94600 leishmania	c 609	8	1.1	1665	16	Q53215	Q53215 mycobacteri
536	8	1.1	660	11	Q9D2E0	Q9D2E0 mus musculu	c 610	8	1.1	1783	5	Q9VX48	Q9VX48 drosophila
537	8	1.1	663	16	Q9KLE2	Q9KLE2 vibrio chol	c 611	8	1.1	1847	5	Q9NKN5	Q9NKN5 leishmania
538	8	1.1	665	11	Q9QXG2	Q9QXG2 mus musculu	c 612	8	1.1	1957	11	Q9JMJ9	Q9JMJ9 mus musculu
539	8	1.1	669	11	Q9JIH4	Q9JIH4 mus musculu	c 613	8	1.1	2083	5	Q60997	Q60997 mus musculu
540	8	1.1	671	10	Q93189	Q93189 arabidopsis	c 614	8	1.1	2090	5	Q9W2T1	Q9W2T1 drosophila
541	8	1.1	671	10	Q91VB7	Q91VB7 arabidopsis	c 615	8	1.1	2157	4	Q9S875	Q9S875 homo sapien
542	8	1.1	671	10	Q94BX7	Q94BX7 arabidopsis	c 616	8	1.1	2157	4	Q96QC6	Q96QC6 homo sapien
543	8	1.1	680	5	Q8SY65	Q8SY65 drosophila	c 617	8	1.1	2487	6	Q9N1T0	Q9N1T0 ornithorhyn
544	8	1.1	680	10	Q9SY66	Q9SY66 arabidopsis	618	8	1.1	3034	12	Q8V1Y4	Q8V1Y4 wheat strea
545	8	1.1	683	10	Q23506	Q23506 arabidopsis	619	8	1.1	3035	12	Q9E1A1	Q9E1A1 wheat strea
546	8	1.1	684	4	Q8TC67	Q8TC67 homo sapien	620	8	1.1	3035	12	Q8V1Y3	Q8V1Y3 wheat strea
547	8	1.1	685	10	Q41382	Q41382 spinacia ol	c 621	8	1.1	3389	4	Q96QU9	Q96QU9 homo sapien
548	8	1.1	690	4	Q8WVZ6	Q8WVZ6 homo sapien	c 622	8	1.1	3508	4	Q96RM4	Q96RM4 homo sapien
549	8	1.1	691	10	Q9FKZ5	Q9FKZ5 arabidopsis	c 623	8	1.1	3616	13	Q9M6V0	Q9M6V0 gallus gall
550	8	1.1	698	10	Q8S1J0	Q8S1J0 oryza sativ	c 624	8	1.1	20	11	Q9QUU8	Q9QUU8 mus sp. c-m
551	8	1.1	699	10	Q8W1K7	Q8W1K7 chlamydomon	c 625	8	1.1	29	2	Q9RC76	Q9RC76 streptococc
552	8	1.1	701	10	Q9S0R8	Q9S0R8 arabidopsis	626	8	1.1	31	2	Q9RC75	Q9RC75 streptococc
553	8	1.1	703	10	Q9LID3	Q9LID3 arabidopsis	c 627	8	1.1	31	15	Q11551	Q11551 human immun
554	8	1.1	709	3	Q60173	Q60173 schizosacch	628	8	1.1	33	2	P71125	P71125 campylobact
555	8	1.1	713	10	Q9MR66	Q9MR66 phaseolus a	c 629	8	1.1	35	2	Q9RC74	Q9RC74 streptococc
556	8	1.1	715	3	Q8TFL3	Q8TFL3 candida gla	c 630	8	1.1	35	4	Q13264	Q13264 homo sapien
557	8	1.1	727	5	Q9VNV3	Q9VNV3 drosophila	c 631	8	1.1	37	4	Q8TDB1	Q8TDB1 homo sapien
558	8	1.1	734	11	Q9PMJ9	Q9PMJ9 mus musculu	c 632	8	1.1	39	2	Q9RC78	Q9RC78 streptococc
559	8	1.1	737	4	Q9BQ39	Q9BQ39 homo sapien	c 633	8	1.1	42	16	Q8X9U0	Q8X9U0 escherichia
560	8	1.1	748	10	Q9SBR6	Q9SBR6 arabidopsis	634	8	1.1	45	16	Q50794	Q50794 borrelia bu
561	8	1.1	748	10	Q8RUW1	Q8RUW1 arabidopsis	c 635	8	1.1	51	16	Q92K67	Q92K67 rhizobium m
562	8	1.1	755	10	Q9FEP21	Q9FEP21 oryza sativ	c 636	8	1.1	54	10	Q9S8X3	Q9S8X3 triticum ae
563	8	1.1	759	13	Q9DF35	Q9DF35 xenopus lae	c 637	8	1.1	54	10	Q9S8X2	Q9S8X2 triticum ae
564	8	1.1	766	10	Q8W0Z5	Q8W0Z5 arabidopsis	c 638	8	1.1	55	2	Q9RC77	Q9RC77 streptococc
565	8	1.1	784	13	Q90YB5	Q90YB5 gallus gall	c 639	8	1.1	58	2	Q9RC79	Q9RC79 streptococc
566	8	1.1	787	10	Q82362	Q82362 arabidopsis	c 640	8	1.1	63	2	Q9AN47	Q9AN47 bradyrhizob
567	8	1.1	797	10	Q93Y39	Q93Y39 arabidopsis	c 641	8	1.1	63	8	Q20444	Q20444 limulus pol
568	8	1.1	800	13	Q9DF36	Q9DF36 xenopus lae	642	8	1.1	63	16	Q8X5E3	Q8X5E3 escherichia
569	8	1.1	802	5	Q9VJ74	Q9VJ74 drosophila	643	8	1.1	65	5	Q9GTY9	Q9GTY9 caenorhabd
570	8	1.1	805	11	Q9VX34	Q9VX34 mus musculu	c 644	8	1.1	65	16	Q99TK2	Q99TK2 staphylococ
571	8	1.1	805	11	Q9CWN5	Q9CWN5 mus musculu	c 645	8	1.1	66	16	Q98PB9	Q98PB9 rhizobium l
572	8	1.1	810	3	Q9PEV7	Q9PEV7 neurospora	646	8	1.1	68	2	Q52624	Q52624 proteus vul
573	8	1.1	810	3	Q9PEV7	Q9PEV7 neurospora	647	8	1.1	68	10	Q9S967	Q9S967 arabidops
574	8	1.1	813	5	Q9V915	Q9V915 drosophila	c 648	8	1.1	69	10	Q9AKI6	Q9AKI6 arabidops
575	8	1.1	821	17	Q8TUM1	Q8TUM1 mechanosarc	c 649	8	1.1	70	12	Q8UXA5	Q8UXA5 hepatitis c
576	8	1.1	832	10	Q9LRY9	Q9LRY9 arabidopsis	c 650	8	1.1	70	12	Q8UXA3	Q8UXA3 hepatitis c
577	8	1.1	830	10	Q9QW11	Q9QW11 mus musculu	c 651	8	1.1	70	12	Q8UX97	Q8UX97 hepatitis c
578	8	1.1	853	16	Q930Y7	Q930Y7 rhizobium m	c 652	8	1.1	70	12	Q8UX94	Q8UX94 hepatitis c
579	8	1.1	879	16	Q9FBJ2	Q9FBJ2 streptomyce	c 653	8	1.1	70	12	Q8UX89	Q8UX89 hepatitis c
580	8	1.1	883	17	Q8TTF9	Q8TTF9 mechanosarc	c 654	8	1.1	70	12	Q8UX86	Q8UX86 hepatitis c
581	8	1.1	926	4	Q74429	Q74429 schizosacch	c 655	8	1.1	70	12	Q8UX84	Q8UX84 hepatitis c
582	8	1.1	936	4	Q95462	Q95462 homo sapien	c 656	8	1.1	70	12	Q8UX83	Q8UX83 hepatitis c
583	8	1.1	926	4	Q9UDR5	Q9UDR5 homo sapien	c 657	8	1.1	70	12	Q8UX82	Q8UX82 hepatitis c
584	8	1.1	934	10	Q94HQ0	Q94HQ0 oryza sativ	c 658	8	1.1	70	12	Q8UX80	Q8UX80 hepatitis c
585	8	1.1	945	5	Q9NKT7	Q9NKT7 leishmania							

C 659	7	1.0	70	12	Q8UX78	Q8UX78 hepatitis c	C 732	7	1.0	128	12	Q68173	Q68173 hepatitis c
C 660	7	1.0	70	12	Q8UX76	Q8UX76 hepatitis c	C 733	7	1.0	129	8	Q9B130	Q9B130 aedes aegypti
C 661	7	1.0	70	12	Q8UX74	Q8UX74 hepatitis c	C 734	7	1.0	129	8	Q9B0U4	Q9B0U4 aedes aegypti
C 662	7	1.0	70	12	Q8UX69	Q8UX69 hepatitis c	C 735	7	1.0	129	8	Q9B0U3	Q9B0U3 aedes aegypti
C 663	7	1.0	70	12	Q8UX67	Q8UX67 hepatitis c	C 736	7	1.0	129	8	Q9B0U2	Q9B0U2 aedes aegypti
C 664	7	1.0	70	12	Q8UX66	Q8UX66 hepatitis c	C 737	7	1.0	129	8	Q9B3H4	Q9B3H4 aedes aegypti
C 665	7	1.0	70	12	Q8UX65	Q8UX65 hepatitis c	C 738	7	1.0	129	8	Q9B3H3	Q9B3H3 aedes aegypti
C 666	7	1.0	71	4	Q9NPy6	Q9NPy6 homo sapien	C 739	7	1.0	129	8	Q9B3H2	Q9B3H2 aedes aegypti
C 667	7	1.0	71	17	Q8TSL0	Q8TSL0 methanosa	C 740	7	1.0	129	8	Q9B3H1	Q9B3H1 aedes aegypti
C 668	7	1.0	71	17	Q8TSL0	Q8TSL0 methanosa	C 741	7	1.0	129	8	Q9B3H0	Q9B3H0 aedes aegypti
C 669	7	1.0	72	6	Q28795	Q28795 layasu taj	C 742	7	1.0	129	8	Q9B3G9	Q9B3G9 aedes aegypti
C 670	7	1.0	72	16	Q9K733	Q9K733 streptomyce	C 743	7	1.0	129	8	Q9B3G8	Q9B3G8 aedes aegypti
C 671	7	1.0	73	5	Q9V1Q3	Q9V1Q3 drosophila	C 744	7	1.0	129	8	Q9B3G7	Q9B3G7 aedes aegypti
C 672	7	1.0	75	2	Q9NLM8	Q9NLM8 leishmania	C 745	7	1.0	129	10	Q9B3XW1	Q9B3XW1 lithospermum
C 673	7	1.0	78	2	Q51924	Q51924 pasteurella	C 746	7	1.0	129	12	Q8BWA0	Q8BWA0 white spot
C 674	7	1.0	80	10	Q04126	Q04126 zea mays (m	C 747	7	1.0	130	8	Q9MDS0	Q9MDS0 aedes aegypti
C 675	7	1.0	83	16	Q8UEM9	Q8UEM9 agrobacteri	C 748	7	1.0	130	8	Q9MDS9	Q9MDS9 aedes aegypti
C 676	7	1.0	85	16	Q9A015	Q9A015 streptococ	C 749	7	1.0	130	16	Q9EC14	Q9EC14 streptomyce
C 677	7	1.0	85	16	Q9BML7	Q9BML7 rhizobium 1	C 750	7	1.0	131	11	Q99L27	Q99L27 mus musculus
C 678	7	1.0	85	16	Q9BML7	Q9BML7 rhizobium 1	C 751	7	1.0	131	15	Q8B389	Q8B389 primate t-1
C 679	7	1.0	86	12	Q9B8X2	Q9B8X2 streptomyce	C 752	7	1.0	131	17	Q8TW52	Q8TW52 methanopyru
C 680	7	1.0	86	12	Q919M4	Q919M4 cullex nigri	C 753	7	1.0	132	6	Q9GMD5	Q9GMD5 macaca fasc
C 681	7	1.0	90	16	Q83111	Q83111 treponema p	C 754	7	1.0	132	16	Q8XD66	Q8XD66 escherichia
C 682	7	1.0	91	13	Q9YH10	Q9YH10 ginglymose	C 755	7	1.0	133	11	Q921B5	Q921B5 mus musculus
C 683	7	1.0	91	16	Q9JX58	Q9JX58 neisseria m	C 756	7	1.0	133	10	Q40447	Q40447 nicotiana g
C 684	7	1.0	91	17	Q8ZKQ1	Q8ZKQ1 pyrobaculum	C 757	7	1.0	134	10	Q884D4	Q884D4 podococcus
C 685	7	1.0	92	16	Q9ZK17	Q9ZK17 helicobacter	C 758	7	1.0	135	2	P70762	P70762 anabaena sp
C 686	7	1.0	93	10	Q9AUN6	Q9AUN6 oryza sativ	C 759	7	1.0	135	5	Q44206	Q44206 megaselia s
C 687	7	1.0	95	5	Q00088	Q00088 drosophila	C 760	7	1.0	135	5	Q9V0G0	Q9V0G0 drosophila
C 688	7	1.0	96	2	Q92914	Q92914 streptomyce	C 761	7	1.0	135	6	Q9GMC4	Q9GMC4 ovine atles
C 689	7	1.0	96	5	Q9VKS2	Q9VKS2 drosophila	C 762	7	1.0	135	15	Q80829	Q80829 human t-cel
C 690	7	1.0	96	10	Q9SDK2	Q9SDK2 oryza sativ	C 763	7	1.0	136	17	Q9HMG2	Q9HMG2 halobacteri
C 691	7	1.0	96	10	Q94HN5	Q94HN5 oryza sativ	C 764	7	1.0	138	11	Q9D3Y0	Q9D3Y0 mus musculus
C 692	7	1.0	98	5	Q23908	Q23908 dictyosteli	C 765	7	1.0	138	15	Q80823	Q80823 human t-cel
C 693	7	1.0	98	5	Q9VCP2	Q9VCP2 drosophila	C 766	7	1.0	139	16	Q8U519	Q8U519 human t-cel
C 694	7	1.0	98	12	Q8QX99	Q8QX99 garlic late	C 767	7	1.0	140	16	Q99RP2	Q99RP2 streptococ
C 695	7	1.0	99	16	Q8YRF9	Q8YRF9 anabaena sp	C 768	7	1.0	141	5	Q967N1	Q967N1 micrarchia ca
C 696	7	1.0	100	8	Q95207	Q95207 lycopersico	C 769	7	1.0	141	6	Q9BER7	Q9BER7 leoparcus p
C 697	7	1.0	100	8	Q95206	Q95206 podophyllum	C 770	7	1.0	142	2	P82552	P82552 streptococ
C 698	7	1.0	104	10	Q9LENT	Q9LENT arabidopsis	C 771	7	1.0	142	4	Q8WM60	Q8WM60 homo sapien
C 699	7	1.0	104	16	Q31480	Q31480 bacillus su	C 772	7	1.0	143	3	Q74666	Q74666 rhizopus or
C 700	7	1.0	104	16	Q8XV14	Q8XV14 ralsstonia s	C 773	7	1.0	143	5	Q18333	Q18333 caenorhabdi
C 701	7	1.0	105	11	Q99J74	Q99J74 mus musculu	C 774	7	1.0	143	6	Q9BBER6	Q9BBER6 panthera on
C 702	7	1.0	105	16	Q99S32	Q99S32 staphylococ	C 775	7	1.0	144	5	Q16707	Q16707 caenorhabdi
C 703	7	1.0	106	2	Q05681	Q05681 mycobacteri	C 776	7	1.0	144	5	Q9GZ13	Q9GZ13 caenorhabdi
C 704	7	1.0	106	10	Q9FTU0	Q9FTU0 oxalis regn	C 777	7	1.0	144	5	Q9RME9	Q9RME9 arabidopsis
C 705	7	1.0	107	5	Q97028	Q97028 dugesia jap	C 778	7	1.0	144	16	Q9PC82	Q9PC82 xylella fas
C 706	7	1.0	107	17	Q96620	Q96620 sulfolobus	C 779	7	1.0	144	16	Q9A7T0	Q9A7T0 caulobacter
C 707	7	1.0	108	8	Q94V73	Q94V73 articia ages	C 780	7	1.0	145	10	P93293	P93293 arabidopsis
C 708	7	1.0	108	8	Q94V73	Q94V73 articia ages	C 781	7	1.0	146	5	Q9XZT7	Q9XZT7 drosophila
C 709	7	1.0	109	5	Q9W486	Q9W486 drosophila	C 782	7	1.0	146	8	Q9TAN9	Q9TAN9 australopit
C 710	7	1.0	109	16	Q9A375	Q9A375 caulobacter	C 783	7	1.0	146	16	Q97NF2	Q97NF2 streptococ
C 711	7	1.0	109	16	Q9A375	Q9A375 caulobacter	C 784	7	1.0	147	10	Q9W7Y1	Q9W7Y1 mus musculu
C 712	7	1.0	109	16	Q933B9	Q933B9 staphylococ	C 785	7	1.0	147	11	Q9D1V3	Q9D1V3 mus musculu
C 713	7	1.0	111	2	Q85208	Q85208 vibrio chol	C 786	7	1.0	148	5	Q96ZAO	Q96ZAO hyalates po
C 714	7	1.0	111	5	Q97024	Q97024 dugesia jap	C 787	7	1.0	148	5	Q8SWM0	Q8SWM0 encephalito
C 715	7	1.0	111	16	Q8YOB4	Q8YOB4 ralsstonia s	C 788	7	1.0	148	6	Q9BEV0	Q9BEV0 talpa alta
C 716	7	1.0	112	3	Q07255	Q07255 saccharomyc	C 789	7	1.0	148	6	Q9BEV0	Q9BEV0 condylura c
C 717	7	1.0	112	17	Q97ZU5	Q97ZU5 sulfolobus	C 790	7	1.0	148	6	Q9BEV0	Q9BEV0 sorex arane
C 718	7	1.0	113	5	Q15720	Q15720 dictyosteli	C 791	7	1.0	148	6	Q9BBER3	Q9BBER3 orycteropu
C 719	7	1.0	113	17	Q96XQ9	Q96XQ9 sulfolobus	C 792	7	1.0	148	6	Q9BBER3	Q9BBER3 caprius ind
C 720	7	1.0	114	10	Q9SFR3	Q9SFR3 oryza sativ	C 793	7	1.0	148	6	Q9BBER8	Q9BBER8 felis silve
C 721	7	1.0	116	17	Q9YDL6	Q9YDL6 aeropyrum p	C 794	7	1.0	148	6	Q9BBER5	Q9BBER5 ursus arcto
C 722	7	1.0	118	5	Q21028	Q21028 caenorhabdi	C 795	7	1.0	148	11	Q9S9M3	Q9S9M3 pedetes cap
C 723	7	1.0	120	5	Q45262	Q45262 caenorhabdi	C 796	7	1.0	148	11	Q9S9M3	Q9S9M3 dipodomys h
C 724	7	1.0	121	16	Q9PEV3	Q9PEV3 xylella fas	C 797	7	1.0	148	11	Q9S9M3	Q9S9M3 cavia teschu
C 725	7	1.0	122	5	Q44205	Q44205 megaselia s	C 798	7	1.0	148	11	Q9S9M3	Q9S9M3 hydrochoeru
C 726	7	1.0	124	10	Q49587	Q49587 arabidopsis	C 799	7	1.0	148	11	Q9S9M3	Q9S9M3 dimomys bra
C 727	7	1.0	124	12	Q9DKM1	Q9DKM1 spodiopetra	C 800	7	1.0	149	5	Q967M8	Q967M8 premobius
C 728	7	1.0	125	3	Q96VH3	Q96VH3 saccharomyc	C 801	7	1.0	149	5	Q96311	Q96311 dendroctonu
C 729	7	1.0	126	10	Q64723	Q64723 arabidopsis	C 802	7	1.0	149	5	Q961Z4	Q961Z4 dendroctonu
C 730	7	1.0	126	16	Q97TQ7	Q97TQ7 clostridium	C 803	7	1.0	149	5	Q961Z0	Q961Z0 hylurgus ml
C 731	7	1.0	127	17	Q9Y9E2	Q9Y9E2 aeropyrum p	C 804	7	1.0	149	11	Q9D7P2	Q9D7P2 mus musculu

C 805	7	1.0	149	12	Q68719	Q68719 hepatitis c	C 878	7	1.0	174	6	Q8MNT5	Q8wnt5 canis lupus
C 806	7	1.0	149	15	Q82393	Q82393 human t-cel	C 879	7	1.0	174	16	Q914R7	Q914r7 pseudomonas
C 807	7	1.0	149	15	Q82395	Q82395 human t-cel	C 880	7	1.0	174	16	Q86519	Q86519 streptomyc
C 808	7	1.0	149	15	Q82397	Q82397 human t-cel	C 881	7	1.0	174	16	Q9R7V6	Q9r7v6 streptomyc
C 809	7	1.0	149	15	Q82399	Q82399 human t-cel	C 882	7	1.0	175	10	Q65088	Q65088 picea maria
C 810	7	1.0	149	15	Q82405	Q82405 human t-cel	C 883	7	1.0	175	11	Q9CYS3	Q9cys3 mus musculu
C 811	7	1.0	149	15	Q82408	Q82408 human t-cel	C 884	7	1.0	175	16	Q8XFS7	Q8xf57 salmonella
C 812	7	1.0	150	13	Q96122	Q96122 hylurdirecto	C 885	7	1.0	176	2	Q9RCU0	Q9rcu0 streptococ
C 813	7	1.0	150	13	Q96122	Q96122 hylurdirecto	C 886	7	1.0	176	14	Q991T7	Q991t7 uncultured
C 814	7	1.0	150	13	Q96124	Q96124 pagrus major	C 887	7	1.0	177	2	Q9KIH6	Q9kih6 rhizobium e
C 815	7	1.0	150	13	Q9PSH8	Q9ps8 xenopus lae	C 888	7	1.0	177	10	Q48838	Q48838 arabidopsis
C 816	7	1.0	150	16	Q91L0P	Q91lp6 streptomyc	C 889	7	1.0	177	10	Q9AV52	Q9av52 oryza sativ
C 817	7	1.0	151	11	Q99P01	Q99p01 mus musculu	C 890	7	1.0	177	10	Q8RW66	Q8rw66 arabidopsis
C 818	7	1.0	151	12	Q12669	Q12669 colobus mon	C 891	7	1.0	177	16	Q8XAS2	Q8xas2 escherichia
C 819	7	1.0	151	15	Q97263	Q97263 human immun	C 892	7	1.0	178	5	Q8SXD7	Q8sxd7 drosophila
C 820	7	1.0	152	5	Q96128	Q96128 dendroctonu	C 893	7	1.0	178	11	Q91X12	Q91x12 cavia porce
C 821	7	1.0	153	5	Q96123	Q96123 hylurdirecto	C 894	7	1.0	178	16	Q9PBT1	Q9pbt1 xyella las
C 822	7	1.0	153	16	Q8YSAB	Q8ysab anabaena sp	C 895	7	1.0	180	5	Q97222	Q97222 plasmodium
C 823	7	1.0	153	17	Q9YFD4	Q9yfd4 aeropyrum p	C 896	7	1.0	180	5	Q9VC28	Q9vc28 drosophila
C 824	7	1.0	154	5	Q9VET5	Q9vet5 drosophila	C 897	7	1.0	180	10	Q9M9M0	Q9m9m0 arabidopsis
C 825	7	1.0	155	5	Q96121	Q96121 hyluronotu	C 898	7	1.0	180	17	Q8U494	Q8u494 pyrococcus
C 826	7	1.0	155	5	Q9VZK6	Q9vzk6 drosophila	C 899	7	1.0	181	5	Q9GZ14	Q9gz14 caenorhabdi
C 827	7	1.0	155	8	Q9GCTM4	Q9gctm4 ochlerotatu	C 900	7	1.0	181	5	Q9WJ33	Q9wj33 drosophila
C 828	7	1.0	155	11	Q99PT6	Q99pt6 mus musculu	C 901	7	1.0	181	5	Q9VWM4	Q9vwm4 drosophila
C 829	7	1.0	155	16	Q97G78	Q97g78 clostridium	C 902	7	1.0	181	10	Q9SBB9	Q9sbb9 arabidopsis
C 830	7	1.0	156	5	Q96127	Q96127 dendroctonu	C 903	7	1.0	181	16	Q9X833	Q9x833 streptomyc
C 831	7	1.0	157	10	Q9AVD9	Q9avd9 nicotiana t	C 904	7	1.0	182	10	Q948P5	Q948p5 oryza sativ
C 832	7	1.0	157	10	Q9AVD2	Q9avd2 nicotiana t	C 905	7	1.0	182	16	Q9KZK9	Q9kzk9 streptomyc
C 833	7	1.0	157	10	Q9AVD1	Q9avd1 nicotiana t	C 906	7	1.0	183	5	Q961Y9	Q961y9 pachycoes
C 834	7	1.0	157	10	Q40238	Q40238 lycopersico	C 907	7	1.0	183	16	Q9L140	Q9l140 streptomyc
C 835	7	1.0	158	5	Q963H5	Q963h5 pityophthor	C 908	7	1.0	184	12	Q89571	Q89571 african swi
C 836	7	1.0	158	5	Q963H4	Q963h4 pityophthor	C 909	7	1.0	184	16	Q91235	Q91235 streptomyc
C 837	7	1.0	158	10	Q9AVD8	Q9avd8 nicotiana t	C 910	7	1.0	186	2	Q56745	Q56745 wolbachia s
C 838	7	1.0	158	10	Q9FOL8	Q9fol8 cyanophora	C 911	7	1.0	186	10	Q9EF14	Q9ef14 arabidopsis
C 839	7	1.0	158	10	Q9ZM37	Q9z37 arabidopsis	C 912	7	1.0	186	10	Q94EP9	Q94ep9 oryza sativ
C 840	7	1.0	159	3	Q12135	Q12135 saccharomyc	C 913	7	1.0	186	10	Q8SA76	Q8sa76 oryza sativ
C 841	7	1.0	159	10	Q65683	Q65683 arabidopsis	C 914	7	1.0	187	17	Q27994	Q27994 archaeoglob
C 842	7	1.0	159	16	Q9RVK2	Q9rvk2 deinococcus	C 915	7	1.0	188	10	Q82214	Q82214 arabidopsis
C 843	7	1.0	160	10	P94060	P94060 arabidopsis	C 916	7	1.0	188	10	Q9SJK7	Q9sjk7 arabidopsis
C 844	7	1.0	160	11	Q35058	Q35058 mus musculu	C 917	7	1.0	188	11	Q9DB11	Q9db11 mus musculu
C 845	7	1.0	161	5	Q963H7	Q963h7 hypochinema	C 918	7	1.0	189	10	Q94U56	Q94u56 oryza sativ
C 846	7	1.0	161	16	Q8ZB19	Q8zb19 yersinia pe	C 919	7	1.0	190	5	Q962A1	Q962a1 hylurgops r
C 847	7	1.0	161	16	Q8Y9R1	Q8y9r1 listeria mo	C 920	7	1.0	191	6	Q9XS09	Q9xs09 capra hircu
C 848	7	1.0	161	17	Q8RTN3	Q8rtn3 methanosarc	C 921	7	1.0	191	4	Q9H3E4	Q9h3e4 homo sapien
C 849	7	1.0	162	5	Q61829	Q61829 caenorhabdi	C 922	7	1.0	191	5	Q46029	Q46029 chironomus
C 850	7	1.0	163	2	Q8VLS9	Q8vls9 wolbachia s	C 923	7	1.0	191	5	Q9W222	Q9w222 drosophila
C 851	7	1.0	163	11	Q54719	Q54719 raltus norv	C 924	7	1.0	191	16	Q9XAE1	Q9xae1 streptomyc
C 852	7	1.0	163	12	Q8VB89	Q8vb89 white spot	C 925	7	1.0	191	17	Q8TZB0	Q8tzb0 pyrobaculum
C 853	7	1.0	164	6	P79361	P79361 ovis aries	C 926	7	1.0	192	10	Q8TZB0	Q8tzb0 methanopyru
C 854	7	1.0	164	10	Q48567	Q48567 euphorbia e	C 927	7	1.0	192	10	Q9FIT9	Q9fit9 arabidopsis
C 855	7	1.0	165	11	Q9CPU3	Q9cpu3 mus musculu	C 928	7	1.0	192	10	Q9FW23	Q9fw23 oryza sativ
C 856	7	1.0	165	11	Q35067	Q35067 mus musculu	C 929	7	1.0	192	12	Q81414	Q81414 hepatitis c
C 857	7	1.0	166	16	Q8XB58	Q8xb58 escherichia	C 930	7	1.0	193	12	Q42623	Q42623 magnaporthe
C 858	7	1.0	166	11	Q35065	Q35065 mus musculu	C 931	7	1.0	193	3	Q42623	Q42623 hepatitis c
C 859	7	1.0	167	10	Q9AVD4	Q9avd4 nicotiana t	C 932	7	1.0	195	5	Q8SV22	Q8sv22 drosophila
C 860	7	1.0	167	10	Q9AVD3	Q9avd3 nicotiana t	C 933	7	1.0	195	5	Q8SV22	Q8sv22 drosophila
C 861	7	1.0	167	10	Q64835	Q64835 arabidopsis	C 934	7	1.0	195	15	Q9DV50	Q9dv50 human immun
C 862	7	1.0	167	16	Q9PCM7	Q9pcm7 xyella las	C 935	7	1.0	196	2	Q48632	Q48632 mycobacteri
C 863	7	1.0	168	5	P91969	P91969 trichinella	C 936	7	1.0	197	2	Q68889	Q68889 lactococcus
C 864	7	1.0	168	11	Q35068	Q35068 mus musculu	C 937	7	1.0	198	5	Q9VPM6	Q9vpm6 drosophila
C 865	7	1.0	168	15	Q9DKR6	Q9dkr6 human t-cel	C 938	7	1.0	199	4	Q96EV2	Q96ev2 homo sapien
C 866	7	1.0	168	16	Q9RV18	Q9rv18 deinococcus	C 939	7	1.0	199	5	Q8T2K7	Q8t2k7 dictyosteli
C 867	7	1.0	168	16	Q98GN6	Q98gn6 rhizobium l	C 940	7	1.0	199	11	Q922C9	Q922c9 mus musculu
C 868	7	1.0	169	5	Q9XUJ8	Q9xuj8 caenorhabdi	C 941	7	1.0	200	5	Q9VEN6	Q9ven6 drosophila
C 869	7	1.0	169	11	Q35061	Q35061 mus musculu	C 942	7	1.0	201	5	Q9VJ78	Q9vj78 drosophila
C 870	7	1.0	170	10	Q9AVD7	Q9avd7 nicotiana t	C 943	7	1.0	201	16	Q9KZK3	Q9kzk3 streptomyc
C 871	7	1.0	170	15	Q70642	Q70642 simian t-ly	C 944	7	1.0	202	10	Q9FEW8	Q9f878 lycopersico
C 872	7	1.0	170	15	Q9W116	Q9w116 human t-cel	C 945	7	1.0	202	10	Q3W877	Q3w877 arabidopsis
C 873	7	1.0	170	15	Q9DKS0	Q9dks0 human t-cel	C 946	7	1.0	202	16	Q8XD16	Q8xd16 escherichia
C 874	7	1.0	170	15	Q73454	Q73454 human t-cel	C 947	7	1.0	202	17	Q8UIP5	Q8ui5 pyrococcus
C 875	7	1.0	170	15	Q85601	Q85601 human t-cel	C 948	7	1.0	202	17	Q8TZ73	Q8tz73 methanopyru
C 876	7	1.0	170	15	Q08838	Q08838 human t-cel	C 949	7	1.0	204	2	Q9AJPI	Q9ajpi streptomyc
C 877	7	1.0	171	12	Q83420	Q83420 pseudorabid	C 950	7	1.0	204	5	Q963H6	Q963h6 pityogenes

```

951 7 1.0 204 5 Q9VP14 Q9VP14 drosophila
952 7 1.0 204 10 Q9LGY5 Q9LGY5 oryza sativ
953 7 1.0 204 16 Q8XEM7 Q8XEM7 escherichia
954 7 1.0 204 16 Q9KZ18 Q9KZ18 streptomyces
955 7 1.0 205 17 Q9YFL1 Q9YFL1 acetylpurum p
956 7 1.0 206 8 Q9B1X9 Q9B1X9 lutzomyia l
957 7 1.0 206 8 Q9B1U7 Q9B1U7 lutzomyia l
958 7 1.0 206 8 Q9B1F4 Q9B1F4 lutzomyia l
959 7 1.0 206 8 Q9B1F3 Q9B1F3 lutzomyia l
960 7 1.0 206 8 Q9B8C1 Q9B8C1 lutzomyia l
961 7 1.0 206 8 Q9B8C0 Q9B8C0 lutzomyia l
962 7 1.0 206 8 Q9B8B9 Q9B8B9 lutzomyia l
963 7 1.0 206 8 Q9B8B8 Q9B8B8 lutzomyia l
964 7 1.0 206 8 Q9B8B7 Q9B8B7 lutzomyia l
965 7 1.0 206 8 Q9B8B6 Q9B8B6 lutzomyia l
966 7 1.0 206 8 Q9B8B5 Q9B8B5 lutzomyia l
967 7 1.0 206 8 Q9B8B4 Q9B8B4 lutzomyia g
968 7 1.0 206 13 Q90802 Q90802 gallus galli
969 7 1.0 206 16 Q98MH4 Q98MH4 rhizobium l
970 7 1.0 207 3 Q01400 Q01400 neurospora
971 7 1.0 208 10 Q9AVX6 Q9AVX6 guillardia
972 7 1.0 208 10 Q9XGJ5 Q9XGJ5 gnetum gnem
973 7 1.0 208 10 Q8S0Z1 Q8S0Z1 oryza sativ
974 7 1.0 209 3 Q13600 Q13600 schizosacch
975 7 1.0 209 4 Q96SH7 Q96SH7 homo sapien
976 7 1.0 209 11 Q9D9B5 Q9D9B5 mus musculu
977 7 1.0 209 16 Q8UP91 Q8UP91 agrobacteri
978 7 1.0 210 2 Q93HY5 Q93HY5 enterococcu
979 7 1.0 210 10 Q01764 Q01764 phytophthor
980 7 1.0 210 11 Q8R206 Q8R206 mus musculu
981 7 1.0 211 5 Q9N4A6 Q9N4A6 caenorhabdi
982 7 1.0 211 13 Q93333 Q93333 struthio ca
983 7 1.0 211 17 Q8ZVD1 Q8ZVD1 pyrobaculum
984 7 1.0 212 15 Q9YXK2 Q9YXK2 human immun
985 7 1.0 212 15 Q9QL60 Q9QL60 human immun
986 7 1.0 213 11 Q9D0H4 Q9D0H4 mus musculu
987 7 1.0 213 12 Q8V4Z0 Q8V4Z0 monkeypox v
988 7 1.0 214 10 Q9SC42 Q9SC42 pieum sativ
989 7 1.0 215 10 Q9LHV7 Q9LHV7 arabidopsis
990 7 1.0 215 17 Q28708 Q28708 archaeoglob
991 7 1.0 216 2 Q8VU76 Q8VU76 lactococcus
992 7 1.0 216 4 Q9BRA8 Q9BRA8 homo sapien
993 7 1.0 216 5 Q9VLD7 Q9VLD7 drosophila
994 7 1.0 216 10 Q41083 Q41083 pinus taeda
995 7 1.0 216 10 Q23595 Q23595 arabidopsis
996 7 1.0 216 16 Q33128 Q33128 mycobacteri
997 7 1.0 217 3 Q9UTZ4 Q9UTZ4 schizosacch
998 7 1.0 217 5 Q9XUA2 Q9XUA2 caenorhabdi
999 7 1.0 217 9 Q80221 Q80221 methanobact
1000 7 1.0 217 10 Q9LPI2 Q9LPI2 arabidopsis

```

## ALIGNMENTS

```

RESULT 1
ID O9GV10 PRELIMINARY; PRT; 546 AA.
AC O9GV10;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Vasa-related protein Povasi (Fragment).
GN POVASI.
OS Ephydactia fluvicollis.
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC Haplosclerida; Spongiillidae; Ephydactia.
OX NCBI_TaxID=31330;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21359115; PubMed=11466525;
RY Nishimura K., Nishimura T., Fujisawa T.;
RT "Universal occurrence of the vasa-related genes among metazoans and
their germline expression in Hydra."

```

```

RL Dev. Genes Evol. 211:299-308 (2001).
DR EMBL: AB047385; BAB13310.1; -.
DR HSSP: Q58083; 1HV8.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD_box.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR002202; HMG-CoA_red.
DR Pfam: PF00270; DEAD; 1.
DR Pfam: PF00271; helicase_C; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HelicC; 1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
DR PROSITE: PS00318; HMG_COA_REDUCTASE_2; UNKNOWN_1.
KM ATP-binding; Helicase.
FT NON TER 1
SQ SEQUENCE 546 AA; 58871 MW; 61845629F637A40 CRC64;

```

```

Alignment Scores:
Pred. No.: 5,32e-19 Length: 546
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.73% Indels: 0
DB: 5 Gaps: 0

```

US-09-714-865-15 (1-2172) x Q9GV10 (1-546)

```

QY 955 ATTCCTATCATCTTGACAGAGATTGATGCTTGCTCAACAGGCTTGCGANG 1014
DB 124 IIEPIRIIEIELEUAIAGIYAGAPLEUIMETALACYSALAGIINIRGILYSEGLYLS 143
QY 1015 ACTGCGCTTTCTCTCTACCA 1035
DB 144 TTTAIAIAIAPHELEULEUPRO 150

```

```

RESULT 2
ID Q8QGD0 PRELIMINARY; PRT; 394 AA.
AC Q8QGD0;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DEAD box RNA helicase Vasa.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21906633; PubMed=11909530;
RY Kraut H., Steinbeisser H., Schwarz H., Nusslein-Volhard C.;
RT "An Evolutionary Conserved Region in the vasa 3' UTR targets RNA
RT translation to the Germ Cells in the Zebrafish."
RL Curr. Biol. 12:454-466 (2002).
DR EMBL: AF479823; AAL87142.1; -.
KM Helicase.
SQ SEQUENCE 394 AA; 43662 MW; 78F9ADF1AA0A324 CRC64;

```

```

Alignment Scores:
Pred. No.: 7.47e-17 Length: 394
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.45% Indels: 0
DB: 13 Gaps: 0

```

US-09-714-865-15 (1-2172) x Q8QGD0 (1-394)

```

QY 1840 AATTGGATCTTCTCTCAATGATGATATGTCATGCAATGGGCTACTGTCGT 1899
DB 286 AATPAAAPLEUPROSETHRIIEASPIGTYVALHISARGILEGLYARGTHRGYARG 305

```



QY 1900 TGTGGATATCTGCC 1914  
 DB 306 CysGlyAsnThrGly 310  
 RESULT 3  
 Q918L8  
 ID Q918L8 PRELIMINARY; PRT; 399 AA.  
 AC Q918L8;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Vasa-like protein (Fragment).  
 OS Danio dangila.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxId=127599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20273912; PubMed=10811828;  
 RA Knaut H., Pelegri F., Bohmann K., Schwarz H., Nusslein-Volhard C.;  
 RT "zebrafish vasa RNA but not its protein is a component of the germ  
 RT plasm and segregates asymmetrically before germline specification.";  
 RL J. Cell Biol. 149:875-888(2000).  
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
 DR EMBL; AF251800; AAF74278.2; -;  
 DR HSP; O58083; IHV8.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR000628; DEAD\_box.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 KM ATP-binding; Helicase; RNA-binding.  
 FT NON\_TER 399 399  
 SQ SEQUENCE 399 AA; 43805 MW; F20B0FA028FA1128 CRC64;  
 Alignment Scores:  
 Pred. No.: 8,58e-16 Length: 399  
 Score: 24.00 Matches: 24  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.31% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-714-865-15 (1-2172) x Q918L8 (1-399)  
 QY 1669 GCACCTTTCTTGTCAAGAAAAATATCAACTACAGTATCGATGGTATCGGAAACAG 1728  
 DB 229 AlathrPhelencysclngluysliiserlthserlthslgslasparlgluGln 248  
 QY 1729 AGAGACGGGAG 1740  
 DB 249 ArgGluArgGlu 252  
 RESULT 4  
 Q90ZF6  
 ID Q90ZF6 PRELIMINARY; PRT; 617 AA.  
 AC Q90ZF6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE VASA.  
 OS Oryzias latipes (Medaka fish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
 OC Belontiiformes; Adrainichthyidae; Oryziinae; Oryziias.  
 OC NCBI\_TaxId=8090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Shinomiya A., Tanaka M., Kobayashi T., Nagahama Y., Hamaguchi S.;  
 RT "The vasa-like gene, olvas, identifies the migration path of  
 RT primordial germ cells during embryonic body formation stage in the  
 RT medaka, Oryzias latipes.";  
 RL Development 42:317-326(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21126920; PubMed=11226275;  
 RA Tanaka M., Kinoshita M., Kobayashi D., Nagahama Y.;  
 RT "Establishment of medaka (Oryzias latipes) transgenic lines with the  
 RT expression of green fluorescent protein fluorescent exclusively in  
 RT germ cells: A useful model to monitor germ cells in a live  
 RT vertebrate.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:2544-2549(2001).  
 DR EMBL; AB063484; BAB61047.1; -;  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR000629; DEAD\_box.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; UNKNOWN\_1.  
 KM ATP-binding; Helicase.  
 SQ SEQUENCE 617 AA; 66737 MW; C1B8A3462760BA50 CRC64;  
 Alignment Scores:  
 Pred. No.: 7,91e-16 Length: 617  
 Score: 24.00 Matches: 24  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.31% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-714-865-15 (1-2172) x Q90ZF6 (1-617)  
 QY 970 GCAGACGAGATTGCTTGTGCTCAAGAGGCTTGGAGAGCTGCGCTTTCTC 1029  
 DB 217 AlaglyArgAspIeuweclacyslaaglntrnglYserGlylysrhrlaalaPhelcu 236  
 QY 1030 CTACCAATTTTG 1041  
 DB 237 LeuProIleleu 240  
 RESULT 5  
 Q80HL6  
 ID Q80HL6 PRELIMINARY; PRT; 621 AA.  
 AC Q80HL6;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Vasa short form.  
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;  
 OC Cichlidae; Oreochromis.  
 OC NCBI\_TaxId=8128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21664110; PubMed=11804791;  
 RA Kobayashi T., Kajitara-Kobayashi H., Nagahama Y.;  
 RT "Two isoforms of vasa homologs in a teleost fish: their differential  
 RT expression during germ cell differentiation.";  
 RL Mech. Dev. 111:167-171(2002).  
 DR EMBL; AB051835; BAB56110.1; -;  
 SQ SEQUENCE 621 AA; 67651 MW; FOA828FDC5308PAA CRC64;  
 Alignment Scores:  
 Pred. No.: 7,9e-16 Length: 621  
 Score: 24.00 Matches: 24  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 3.31% Indels: 0
DB: 13 Gaps: 0

US-09-714-865-15 (1-2172) x Q8OHL6 (1-621)

OY 970 GCAGACAGATTGATGGCTTGTGCTCAAAACAGGCTCGGGAAGACTGCGGCTTTTCTC 1029
DB 219 AAGGTAAGAGpLeuMeCaLaCySaLaGInThrGlySerGlyLVsThraLaAlaPheLeu 238
OY 1030 CTACCAATTG 1041
DB 239 LeuProlLeu 242

RESULT 6
Q9DEG3 PRELIMINARY; PRT; 645 AA.
ID Q9DEG3
AC Q9DEG3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Vasa.
GN VAS.
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percormorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20544961; PubMed=11091081;
RA Kobayashi T., Kajitani-Kobayashi H., Nagahama Y.;
RT "Differential expression of vasa homologue gene in the germ cells
RT during oogenesis and spermatogenesis in a teleost fish, tilapia,
RT Oreochromis niloticus."
RL Mech. Dev. 99:139-142(2000).
DR EMBL; AB032467; BAB19807.1; -.
DR HSSP; Q58083; IHV8.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR000362; Fumarate_lyase.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR00685; RUBISCO_large.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXdc; 1.
DR SMART; SM00490; HELICG; 1.
DR PROSITE; PS00039; DEND_ATP_HELICASE; 1.
DR PROSITE; PS00163; FUMARATE_LYASES; UNKNOWN_1.
DR PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
DR ATP-binding; Helicase.
SQ SEQUENCE 645 AA; 70072 MW; 4C945536A9BF21DA CRC64;

Alignment Scores:
Pred. No.: 7,846-16 Length: 645
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.31% Indels: 0
DB: 13 Gaps: 0

US-09-714-865-15 (1-2172) x Q9DEG3 (1-645)

OY 970 GCAGACAGATTGATGGCTTGTGCTCAAAACAGGCTCGGGAAGACTGCGGCTTTTCTC 1029
DB 219 AAGGTAAGAGpLeuMeCaLaCySaLaGInThrGlySerGlyLVsThraLaAlaPheLeu 262
OY 1030 CTACCAATTG 1041
DB 263 LeuProlLeu 266

RESULT 7
Q9PT10

```

ID	Q9PT10	PRELIMINARY;	PRT;	647 AA.
AC	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
DT	01-MAY-2002 (TREMBlrel. 20, Last annotation update)			
DE	Vasa.			
GN	Vas.			
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
OC	Poeciliaceae; Cyprinidae; Salmoniformes; Salmonidae; Oncorhynchus.			
CC	NCBI_TaxID=8022;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=OVARY;			
RA	Yoshitake G., Saketani S., Tomiwa H., Takeuchi T.;			
RT	"Cloning and characterization of a vasa-like gene in rainbow trout and			
RT	its expression in the germ cell lineage."			
RL	Mol. Reprod. Dev. 55:364-371(2000).			
CC	-1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.			
DR	EMBL; AB032566; BAA88059.1; -			
DR	HSSP; Q58083; 1HV8.			
DR	InterPro; IPR001410; DEAD.			
DR	InterPro; IPR000629; DEAD_box.			
DR	InterPro; IPR01650; Helicase_C.			
DR	Pfam; PF00270; DEAD. 1.			
DR	Pfam; PF00271; helicase_C. 1.			
DR	SMART; SM00487; DEXDC; 1.			
DR	SMART; SM00480; HELIC_C; 1.			
DR	PROSITE; PS00039; DEAD_ATP_HELICASE; 1.			
DR	ATP-binding; Helicase; RNA-binding.			
KW	SEQUENCE 647 AA; 68525 MW; 88F25AE7A68F51C2 CRC64;			
SO				

  

Alignment Scores:				
Pred. No.:	7.84e-16	Length:	647	
Score:	24.00	Matches:	24	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	3.31%	Indels:	0	
DB:	13	Gaps:	0	

  

US-09-714-865-15 (1-2172) x Q9PT10 (1-647)				
QY	970 GCAGGACGAGATTGATGCGCTTGCTCAACAGAGGCTCGGAGACTGCGGCTTTTCTC	10239		
Db	247 ALaelytArqspuenuMetalaCyalaGlnthrclyserglystnAlaAlaPheleu	266		
QY	1030 CTACCAATTG 1041			
Db	267 LeuPrlleleu 270			

  

RESULT 8				
Q91372	PRELIMINARY;	PRT;	700 AA.	
AC	Q91372;			
DT	01-NOV-1996 (TREMBlrel. 01, Created)			
DT	01-JUN-1998 (TREMBlrel. 06, Last sequence update)			
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)			
DE	DEAD box protein.			
GN	VLI1.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
CC	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=94200507; PubMed=8150200;			
RT	Komiyama T., Itoh K., Ikenishi K., Furusawa M.;			
RT	"Isolation and characterization of a novel gene of the DEAD box			
RT	protein family which is specifically expressed in germ cells of			
RT	Xenopus laevis."			

RL Dev. Biol. 162:354-363(1994).  
RN (12)  
RP SEQUENCE FROM N.A.  
RA Komiyama T.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
DR EMBL; AF046043; AAC03114.1; -  
DR HSSP; Q58083; 1HV8.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR000629; DEAD box.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00270; DEAD; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00490; HELIC; 1.  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
DR PROSITE; PS00157; RUBISCO\_LARGE; UNKNOWN\_1.  
KW ATP-binding; Helicase; RNA-binding.  
SQ SEQUENCE 700 AA; 78240 MW; C504ECA38EFB0B7E CRC64;

Alignment Scores:  
Pred. No.: 7,72e-16 Length: 700  
Score: 24.00 Matches: 24  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.31% Indels: 0  
Gaps: 0  
DB: 13

US-09-714-865-15 (1-2172) x Q91372 (1-700)

QY 970 GCAGACGAGATTGCTGCTGCTCAACAGGCTCTGGAGAGACTGCGGCTTTCTC 1029  
|||  
Db 310 AAGGlyArgAspLeuMetLacCysAlaGlnThrGlySerClyLysThrAlaAlaPheLeu 329  
|||

QY 1030 CTACCAATTTTG 1041  
|||  
Db 330 LeuProIleLeu 333  
|||

RESULT 9

ID Q42378 PRELIMINARY; PRT; 700 AA.  
AC Q42378;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE RNA helicase (DEAD box).  
GN VASA OR VLG.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_Taxid=7955;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98025484; PubMed=9376327;  
RA Olsen L.C., Aasland R., Fjose A.;  
RT "A vasa-like gene in Zebrafish identifies putative primordial germ  
RT cells.";  
RL Mech. Dev. 66:95-105(1997).  
CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
DR EMBL; Y12007; CAA72735.1; -  
DR HSSP; Q58083; 1HV8.  
DR ZFIN; ZDB-GENE-990415-272; vasa.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR000629; DEAD box.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00270; DEAD; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
KW ATP-binding; Helicase; RNA-binding.

SQ SEQUENCE 700 AA; 75312 MW; 726B56A90D6C2DB3 CRC64;

Alignment Scores:  
Pred. No.: 7,72e-16 Length: 700  
Score: 24.00 Matches: 24  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.31% Indels: 0  
Gaps: 0  
DB: 13

US-09-714-865-15 (1-2172) x Q42378 (1-700)

QY 970 GCAGACGAGATTGCTGCTGCTCAACAGGCTCTGGAGAGACTGCGGCTTTCTC 1029  
|||  
Db 297 AAGGlyArgAspLeuMetLacCysAlaGlnThrGlySerClyLysThrAlaAlaPheLeu 316  
|||

QY 1030 CTACCAATTTTG 1041  
|||  
Db 317 LeuProIleLeu 320  
|||

RESULT 10

ID Q8QG8 PRELIMINARY; PRT; 715 AA.  
AC Q8QG8;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Vasa-like protein.  
GN VASA.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_Taxid=7955;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX STRAIN=AB;  
RA Barfai R., Orban L.;  
RT "Characterization of the genomic locus encoding vasa protein in  
RT zebrafish.";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF461759; AAL89410.1; -  
SQ SEQUENCE 715 AA; 76795 MW; 602331B3700203AA CRC64;

Alignment Scores:  
Pred. No.: 7,69e-16 Length: 715  
Score: 24.00 Matches: 24  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.31% Indels: 0  
Gaps: 0  
DB: 13

US-09-714-865-15 (1-2172) x Q8QG8 (1-715)

QY 970 GCAGACGAGATTGCTGCTGCTCAACAGGCTCTGGAGAGACTGCGGCTTTCTC 1029  
|||  
Db 312 AAGGlyArgAspLeuMetLacCysAlaGlnThrGlySerClyLysThrAlaAlaPheLeu 331  
|||

QY 1030 CTACCAATTTTG 1041  
|||  
Db 332 LeuProIleLeu 335  
|||

RESULT 11

ID Q8QF0 PRELIMINARY; PRT; 715 AA.  
AC Q8QF0;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Vasa-like protein.  
GN VLG.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Vatlund A., Olsen L.C.;  
RT "Establishment of transgenic zebrafish lines expressing green  
fluorescent protein in their germ cells."  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ111625; CAC84069.1; -  
SQ SEQUENCE 715 AA; 76880 MW; 3CE4E03F02073BA4 CRC64;  
  
Alignment Scores:  
Pred. No.: 7.69e-16 Length: 715  
Score: 24.00 Matches: 24  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.31% Indels: 0  
DB: 13 Gaps: 0  
  
US-09-714-865-15 (1-2172) x Q8QF00 (1-715)  
QY 970 GCAGACGAGATTGATGCTGTCTCAACAGGCTTCGGAGAGCTGCGCTTTTCTC 1029  
DB 312 AAGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyThrAlaAlaPheLeu 331  
QY 1030 CTACCAATTTTG 1041  
DB 332 LeuProIleLeu 335  
  
RESULT 12  
ID 042107 PRELIMINARY; PRT; 716 AA.  
AC 042107;  
DT 01-JUN-1998 (TrEMBLrel. 05, Created)  
DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Vasa.  
GN VASA.  
OS Brachydanio rerio (zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=97417583; PubMed=9272956;  
RA Yoon C., Kawakami K., Hopkins N.;  
RT "zebrafish vasa homologue RNA is localized to the cleavage planes of  
2- and 4-cell-stage embryos and is expressed in the primordial germ  
cells."  
RL Development 124:3157-3166(1997).  
CC - SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
DR EMBL: AB005147; BAA22535.1; -  
DR HSSP: Q58083; 1HV8.  
DR ZFIN: ZDB-GENE-990415-272; vasa.  
DR InterPro: IPR001410; DEAD.  
DR InterPro: IPR000628; DEAD\_box.  
DR InterPro: IPR001650; Helicase\_C.  
DR Pfam: PF00271; DEAD\_1.  
DR Pfam: PF00271; helicase\_C; 1.  
DR SMART: SM00487; DEKDC; 1.  
DR SMART: SM00490; HELIC\_C; 1.  
DR PROSITE: PS00039; DEAD\_ATP\_HELICASE; 1.  
KW ATP-binding; Helicase; RNA-binding.  
SQ SEQUENCE 716 AA; 77004 MW; BB74923B95C7AC5A CRC64;  
  
Alignment Scores:  
Pred. No.: 7.69e-16 Length: 716  
Score: 24.00 Matches: 24  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.31% Indels: 0

DB: 13 Gaps: 0  
US-09-714-865-15 (1-2172) x 042107 (1-716)  
QY 970 GCAGACGAGATTGATGCTGTCTCAACAGGCTTCGGAGAGCTGCGCTTTTCTC 1029  
DB 312 AAGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyThrAlaAlaPheLeu 332  
QY 1030 CTACCAATTTTG 1041  
DB 333 LeuProIleLeu 336  
  
RESULT 13  
ID 080GC9 PRELIMINARY; PRT; 400 AA.  
AC 080GC9;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE DEAD box RNA helicase Vasa.  
OS Melanotaenia fluviatilis.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Atheriniformes; Melanotaenioidae; Melanotaeniidae; Melanotaenia.  
OX NCBI\_TaxID=120844;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21906633; PubMed=11909530;  
RA Kraut H., Steinbeisser H., Schwarz H., Nusslein-Volhard C.;  
RT "An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA  
Translation to the Germ Cells in the Zebrafish."  
RL Curr. Biol. 12:454-466(2002).  
DR EMBL: AF479824; AAL87143.1; -  
KW Helicase.  
SQ SEQUENCE 400 AA; 43797 MW; 5BDEF82585AC9059 CRC64;  
  
Alignment Scores:  
Pred. No.: 9.86e-15 Length: 400  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.18% Indels: 0  
DB: 13 Gaps: 0  
  
US-09-714-865-15 (1-2172) x 080GC9 (1-400)  
QY 1861 ATTGATGAATATGTCATGGAATGGCGTACTGCTGTGGAGATACGGCAGACA 1920  
DB 293 IleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 312  
QY 1921 ATTTCCTTT 1929  
DB 313 IleSerPhe 315  
  
RESULT 14  
ID 090MS9 PRELIMINARY; PRT; 658 AA.  
AC 090MS9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE DBY protein (DEAD (aspartate-glutamate-alanine-aspartate) box  
polypeptide, Y chromosome).  
DE polyptide, Y chromosome).  
GN DBY OR DBY.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=C57BL/6; TISSUE=TESTIS;  
MEDLINE=98409500; PubMed=9736773;  
RX

RA Mazeyrat S., Saut N., Sargent C., Grimmond S., Longepied G.,  
RA Ehrmann I.E., Ellis P.S., Greenfield A., Affara N.A., Mitchell M.J.,  
RT "The mouse Y chromosome interval necessary for spermatogonial  
RT proliferation is gene dense with syntenic homology to the human AZFa  
RT region." Genet. 7:1713-1724(1998).  
RL Hum. Mol. Genet. 7:1713-1724(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
DR EMBL; A007376; CA07483.1; -;  
DR EMBL; BC021453; AA021453.1; -;  
DR HSSP; OS8083; 1HV8.  
DR MGD; MG1:1349406; Dby.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR000629; DEAD\_box.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00270; DEAD; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
KW ATP-binding; Helicase; RNA-binding.  
SQ SEQUENCE 658 AA; 73427 MW; C64668326B2C3BB9 CRC64;

## Alignment Scores:

Pred. No.: 1.03e-13 Length: 658  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.04% Indels: 0  
DB: 11 Gaps: 0

US-09-714-865-15 (1-2172) x Q9QWS9 (1-658)

QY 976 CGAGATTGATGCTTGTGCTCAACAGAGGCTGGAGACTCGGCTTTCTCTACCA 1035  
DB 217 ArgAspLeuMetLacysaInGlnGlySerGlyThrAlaAlaPheLeuPro 236  
1036 ATTTTG 1041  
237 IleLeu 238

## RESULT 15

Q8QGC8 PRELIMINARY; PRT; 396 AA.  
ID Q8QGC8  
AC Q8QGC8;  
DT 01-JUN-2002 (T-EMBLrel. 21, Created)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE DEAD box RNA helicase Vasa.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21906633; PubMed=11909530;  
RA Knaut H., Steinbeisler H., Schwarz H., Nusslein-Volhard C.;  
RT "An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA  
RT Translation to the Germ Cells in the Zebrafish.";  
RL Curr. Biol. 12:454-466(2002).  
DR EMBL; AF479825; AAU87144.1; -;  
KW Helicase.  
SQ SEQUENCE 396 AA; 43087 MW; FFD9B3FDED17BA0 CRC64;

## Alignment Scores:

Pred. No.: 1.31e-12 Length: 396  
Score: 21.00 Matches: 21  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.90% Indels: 0  
DB: 13 Gaps: 0

US-09-714-865-15 (1-2172) x Q8QGC8 (1-396)

QY 1699 ACTACAAGTATCCATGCTGATCGGAAACAGAGACGGGACCAAGCTTTGGAGATTTT 1758  
DB 239 ThrThrSerIleHisGlyAspArgGlnGlnArgGlnAlaLeuGlyAspPhe 258  
QY 1759 CGC 1761  
DB 259 Arg 259

Search completed: June 10, 2003, 17:08:43  
Job time : 115 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:45:36 ; Search time 21.5 Seconds  
(without alignments)  
8380.147 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 724  
Sequence: 1 atgggggagatgaagattggga.....tagatgatgagcatgggat 2172

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Listing first 1000 summaries

Command line parameters:  
-MODE=frame+n2p.model -DEV=xlp  
-O=/cgn2.1/USPRO.spool/US09714865/runat.05062003.11832.26571/app.query.fasta.1.2311  
-DB=SwissProt\_40 -QPMT=fastran -SUFFIX=oligo.rsp -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -STRAT=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=1000 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09714865 @CGN 1.1 23 @runat.05062003.11832.26571 -NCPU=6 -ICPU=3  
-NO MAP -LARGEDUTERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONCLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	61.1	84.4	724	1	DDX4 HUMAN
2	82	11.3	702	1	DDX4 MOUSE
3	82	11.3	713	1	DDX4 RAT
4	23	3.2	661	1	VASA DROME
5	22	3.0	661	1	PLI0 MOUSE
6	22	3.0	661	1	DDX3 HUMAN
7	22	3.0	661	1	DDX3 MOUSE
8	22	3.0	697	1	AN3 XENLA
9	18	2.5	660	1	DDX4 HUMAN
10	18	2.5	660	1	GLH1 CAEEL
11	17	2.3	617	1	DBP1 YEAST
12	17	2.3	720	1	GLH3 CAEEL
13	16	2.2	245	1	DEA2 MOUSE
14	16	2.2	604	1	DED1 YEAST
15	16	2.2	636	1	DED1 SCHPO
16	15	2.1	546	1	DBP2 YEAST
17	15	2.1	550	1	DBP2 SCHPO
18	15	2.1	568	1	HE64 TRYBB

19	15	2.1	575	1	RM62 DROME	P19109 drosophila
20	15	2.1	617	1	DB10 NIOSY	P46942 nicotiana s
21	15	2.1	604	1	DDX5 HUMAN	P17844 homo sapien
22	15	2.1	614	1	DDX5 MOUSE	O6156 mus musculu
23	15	2.1	650	1	DD17 HUMAN	O92841 homo sapien
24	14	1.9	454	1	RHIE ECOLI	P25888 escherichia
25	12	1.7	548	1	SRMB ECOLI	P21507 escherichia
26	12	1.7	478	1	YBX6 SCHPO	Q10202 schizosacch
27	11	1.5	523	1	DBPA ECOLI	P21693 escherichia
28	11	1.5	523	1	DBP3 YEAST	P20447 saccharomyc
29	11	1.5	543	1	RRP3 YEAST	P38712 saccharomyc
30	11	1.5	601	1	DEAD BUCAL	P57453 saccharomyc
31	11	1.5	613	1	DEAD HAEIN	P57453 buchnera ap
32	11	1.5	628	1	DEAD ECOLI	P44586 haemophilus
33	11	1.5	642	1	DEAD KLEPN	P23304 escherichia
34	11	1.5	1156	1	GLH4 CAEEL	P33906 klebsiella
35	10	1.4	418	1	RHUB HAEIN	O76743 caenorhabdi
36	10	1.4	420	1	RHUB ECOLI	P44922 haemophilus
37	10	1.4	520	1	RHUB SALTY	P24229 escherichia
38	10	1.4	424	1	EXP9 STRPN	P40863 salmonella
39	10	1.4	563	1	DEAD MYCTU	P35599 streptococc
40	10	1.4	564	1	ROK1 YEAST	Q11039 mycobacteri
41	10	1.4	619	1	ABS DROME	P45818 saccharomyc
42	10	1.4	622	1	ABS HUMAN	O9V3C0 drosophila
43	9	1.3	168	1	Y168 ADE02	O9U9V9 homo sapien
44	9	1.3	431	1	DBP8 YEAST	P03292 human adeno
45	9	1.3	437	1	TIG1 DROME	P38719 saccharomyc
46	9	1.2	439	1	DDX7 HUMAN	P47980 drosophila
47	9	1.2	439	1	SRMB HAEIN	Q15320 homo sapien
48	9	1.2	453	1	EZY1 CHLRE	P44701 haemophilus
49	9	1.2	481	1	YAB8 SCHPO	Q08354 chlamydomon
50	9	1.3	508	1	EGRI RAT	O09775 schizosacch
51	9	1.3	543	1	EGRI HUMAN	P08154 rattus norv
52	9	1.3	579	1	ESR1 SPRAU	P18146 homo sapien
53	9	1.2	606	1	SPB4 YEAST	O9PVZ9 sparus aura
54	9	1.2	663	1	DUS8 MOUSE	P2808 saccharomyc
55	9	1.2	752	1	DRS1 YEAST	O09112 mus musculu
56	9	1.2	754	1	YAJ3 SCHPO	P32922 saccharomyc
57	9	1.3	788	1	SP1 RAT	O09903 schizosacch
58	9	1.2	848	1	YAA7 SCHPO	Q01714 rattus norv
59	9	1.3	958	1	TRH DROME	O09719 schizosacch
60	9	1.2	995	1	DBPA YEAST	Q24119 drosophila
61	9	1.2	1193	1	DP3A XYLFA	Q12389 saccharomyc
62	9	1.1	115	1	KV2A HUMAN	O9PQ44 xylella fas
63	8	1.1	143	1	RNH HELPY	P01614 homo sapien
64	8	1.1	265	1	UCR1 SOLTU	P56120 heliobacte
65	8	1.1	267	1	COLI FIG	P37841 solinum tub
66	8	1.1	271	1	PDX4 HUMAN	P01192 s corticotr
67	8	1.1	274	1	PDX4 BOVIN	Q13162 homo sapien
68	8	1.1	274	1	PDX4 MOUSE	O9B412 bos taurus
69	8	1.1	297	1	EZY3 CHLRE	O08807 mus musculu
70	8	1.1	320	1	CH36 CERCA	O08356 chlamydomon
71	8	1.1	327	1	MAUR STRPN	P17110 ceretifis c
72	8	1.1	367	1	H669 METUA	O08511 streptococc
73	8	1.1	367	1	HK61 HUMAN	O58083 methanococc
74	8	1.1	416	1	NAP5 HUMAN	P78426 homo sapien
75	8	1.1	460	1	DB80 DROME	O14513 homo sapien
76	8	1.1	464	1	IFB ECALA	O61305 drosophila
77	8	1.1	469	1	DPD2 BOVIN	O04948 biotrichosco
78	8	1.1	478	1	DD19 MOUSE	P49004 bos taurus
79	8	1.1	479	1	DBPA BACSU	O6155 mus musculu
80	8	1.1	479	1	DD19 HUMAN	P42305 bacillus su
81	8	1.1	482	1	DBP5 YEAST	O9UMR2 homo sapien
82	8	1.1	489	1	YRN21 CAEEL	P20442 saccharomyc
83	8	1.1	503	1	YB66 SCHPO	P34580 caenorhabdi
84	8	1.1	505	1	HAS1 YEAST	O09747 schizosacch
85	8	1.1	518	1	FUS MOUSE	O03532 saccharomyc
86	8	1.1	521	1	DB45 DROME	O07868 drosophila
87	8	1.1	526	1	FUS HUMAN	P35637 homo sapien
88	8	1.1	546	1	NOH6 MOUSE	O9D6r4 mus musculu
89	8	1.1	547	1	NOH6 HUMAN	O9NY93 homo sapien
90	8	1.1	564	1	ZYX MOUSE	O62523 mus musculu
91	8	1.1	569	1	KICJ MOUSE	P02535 mus musculu

92	8	1.1	572	1	ZYX_HUMAN	015942	homo sapien	165	7	1.0	185	1	YWLD_BACSU	P39154	bacillus s
93	8	1.1	578	1	YAK2_SCHPO	009916	schizosacch	166	7	1.0	186	1	DNA1_WOLSP	P35907	wolbachia s
94	8	1.1	585	1	ESR1_ORENT	099133	oreochromis	167	7	1.0	189	1	YD48_AQUAE	O67363	aquifex aeo
95	8	1.1	588	1	PR28_YEAST	P23394	saccharomyc	168	7	1.0	190	1	BCT7_SHEEP	P50415	ovis aries
96	8	1.1	617	1	41_BOVIN	091179	bos taurus	169	7	1.0	192	1	PABA_STRL1	P27627	streptomyces
97	8	1.1	633	1	NAH2_YEAST	004121	saccharomyc	170	7	1.0	192	1	PTH_AQUAE	O66677	aquifex aeo
98	8	1.1	663	1	PIT_DROME	094551	drosophila	171	7	1.0	194	1	AX2A_PHAUV	P32293	phaeococcus a
99	8	1.1	670	1	DD18_HUMAN	09nvpi	homo sapien	172	7	1.0	194	1	VATE_HALVO	O48329	halobacteri
100	8	1.1	715	1	DD21_HUMAN	09nvpi	homo sapien	173	7	1.0	202	1	Y1BF_ECOLI	P32105	escherichia
101	8	1.1	742	1	DBP7_YEAST	P36120	saccharomyc	174	7	1.0	205	1	YK07_YEAST	P36661	saccharomyc
102	8	1.1	752	1	PHSM_STREP	P29849	streptococc	175	7	1.0	206	1	UCR1_CHLRE	P49728	chlamydomon
103	8	1.1	770	1	DBP4_YEAST	P20448	saccharomyc	176	7	1.0	209	1	Y0J1_ECOLI	P76657	escherichia
104	8	1.1	773	1	MAK5_YEAST	P38112	saccharomyc	177	7	1.0	210	1	YG21_YEAST	P53251	saccharomyc
105	8	1.1	801	1	41_XENLA	P14434	xenopus lae	178	7	1.0	221	1	IER2_MOUSE	P17950	mus muscucu
106	8	1.1	805	1	E2F_DROME	027368	drosophila	179	7	1.0	223	1	TONB_XANCP	O34761	xanthomonas
107	8	1.1	823	1	NUC1_NEUCR	P20824	neutrospora	180	7	1.0	228	1	UREF_YEREN	P42870	yersinia en
108	8	1.1	851	1	DD21_MOUSE	09j1k5	mus muscucu	181	7	1.0	228	1	UREF_YERPE	O9217	yersinia pe
109	8	1.1	854	1	VG12_BPEZA	P07537	bacteriopho	182	7	1.0	228	1	UREF_YERPS	P52318	yersinia ps
110	8	1.1	857	1	DD24_MOUSE	09esvo	mus muscucu	183	7	1.0	228	1	VIF_OMVS	P16902	ovine lenti
111	8	1.1	858	1	41_MOUSE	P48193	mus muscucu	184	7	1.0	229	1	RAS2_NEUCR	O01387	neutrospora
112	8	1.1	859	1	DD24_HUMAN	09g277	homo sapien	185	7	1.0	229	1	VIF_GAEVC	P33462	caprine art
113	8	1.1	863	1	SYV_ARCFU	028059	archaeoglob	186	7	1.0	230	1	VIF_VILV	P03403	visna lenti
114	8	1.1	875	1	DD10_HUMAN	013206	homo sapien	187	7	1.0	230	1	VIF_VILV1	P23430	visna lenti
115	8	1.1	971	1	SECS_YEAST	P89102	saccharomyc	188	7	1.0	232	1	YM29_MYCLE	O9c5b9	mycobacteri
116	8	1.1	1043	1	CHS2_PABR	092444	paracoccidi	189	7	1.0	236	1	TL17_ARATH	P81760	arabidopsi
117	8	1.1	1235	1	NFT5_MOUSE	09wv30	mus muscucu	190	7	1.0	237	1	ATP6_PETMA	O35538	petromyzon
118	8	1.1	1295	1	GLP1_CAEL	P13508	caenorhabdi	191	7	1.0	238	1	ATP1_CHLRE	O63075	chlamydomon
119	8	1.1	1324	1	SALA_HUMAN	09nec2	homo sapien	192	7	1.0	239	1	G1DB_CLOAB	O97cW4	clostridium
120	8	1.1	1531	1	NFT5_HUMAN	04916	homo sapien	193	7	1.0	240	1	DLX4_MOUSE	P70436	mus muscucu
121	8	1.1	2142	1	BAT2_HUMAN	P48634	homo sapien	194	7	1.0	243	1	PLT_HUMAN	O02083	homo sapien
122	8	1.1	2175	1	HMCU_DROME	P10180	drosophila	195	7	1.0	247	1	ATP1_CHLUV	P56295	chlorella v
123	7	1.0	52	1	ATP8_PENMO	09mgd7	penaeus mon	196	7	1.0	247	1	MCT1_HUMAN	P23346	homo sapien
124	7	1.0	64	1	RL28_MYCCE	P47665	mycoplasma	197	7	1.0	247	1	MCT1_MACFA	P56435	macaca fasc
125	7	1.0	73	1	RL35_MYCLE	09cc21	mycobacteri	198	7	1.0	247	1	MCT1_PAPHA	P52195	papio hamad
126	7	1.0	91	1	YC40_ODOSI	P49535	odontella s	199	7	1.0	247	1	MOC_HUMAN	O16653	homo sapien
127	7	1.0	91	1	RS7_APEPP	09zeus	apple proli	200	7	1.0	248	1	GR1_RAT	O06605	rattus norv
128	7	1.0	91	1	VE5_HPVI3	002267	human papil	201	7	1.0	250	1	ETRA_SHEON	P46148	shevansella
129	7	1.0	96	1	CLV3_ARATH	09xt04	arabidopsi	202	7	1.0	251	1	YMW5_CAEEL	P34507	caenorhabdi
130	7	1.0	104	1	Y164_ARCFU	092815	archaeoglob	203	7	1.0	253	1	LEP4_VIRBCH	P27717	vibrio chol
131	7	1.0	110	1	GON2_SUNMU	097686	suncus murti	204	7	1.0	254	1	UL79_HSVJ	P52470	human herpe
132	7	1.0	112	1	SY27_HUMAN	09y4x3	homo sapien	205	7	1.0	256	1	PXBA_CANBO	O00316	candida boi
133	7	1.0	114	1	GON2_TUPGB	095336	tupaia glis	206	7	1.0	256	1	PXBB_CANBO	O00317	candida boi
134	7	1.0	116	1	YJ57_YEAST	P47132	saccharomyc	207	7	1.0	257	1	ECB8_MYCLE	O07137	mycobacteri
135	7	1.0	117	1	NU3M_PROMI	037625	proctotheca	208	7	1.0	257	1	NT6A_HUMAN	P34132	homo sapien
136	7	1.0	120	1	GON2_HUMAN	043555	homo sapien	209	7	1.0	257	1	NT6B_HUMAN	P34133	homo sapien
137	7	1.0	120	1	R18E_THRAC	09h106	thermoplasma	210	7	1.0	258	1	YC10_METUA	P34133	homo sapien
138	7	1.0	133	1	FOLB_MYCLE	069529	mycobacteri	211	7	1.0	259	1	EMD_MOUSE	O58607	methanococc
139	7	1.0	133	1	POB8_YEAST	P38208	saccharomyc	212	7	1.0	259	1	MSRA_LACSA	O08579	mus muscucu
140	7	1.0	133	1	RS9_DEIRA	09rxv0	delnoccocus	213	7	1.0	262	1	GDAL_WHEAT	O986c2	lactuca satv
141	7	1.0	147	1	RNL4_HUMAN	P34096	homo sapien	214	7	1.0	266	1	ARC4_PHAUV	P04721	trititum ae
142	7	1.0	147	1	RNL4_PIG	P15468	sus scrofa	215	7	1.0	267	1	MM07_RAT	O43629	phaeoculus v
143	7	1.0	147	1	RNL4_PIG	055004	rattus norv	216	7	1.0	268	1	EP34_HOMVA	P50280	rattus norv
144	7	1.0	148	1	RNL4_MOUSE	09j1h1	mus muscucu	217	7	1.0	273	1	YARZ_ECOLI	P16768	human cytom
145	7	1.0	150	1	ARQO_BUCAI	P57479	buchnera ap	218	7	1.0	273	1	YFJO_ECOLI	P77206	escherichia
146	7	1.0	154	1	SM20_SCHMA	P15845	schistosoma	219	7	1.0	274	1	GLPE_BACSU	P52132	escherichia
147	7	1.0	155	1	RS7_MYCLE	092801	chlamydia m	220	7	1.0	274	1	STCO_EMENT	P18156	baecillus su
148	7	1.0	155	1	RS7_MYCCE	P30764	mycobacteri	221	7	1.0	279	1	IVEN_MICLU	O00713	emerichella
149	7	1.0	155	1	RS7_MYCCE	P41193	mycobacteri	222	7	1.0	279	1	CYSW_STNPT	P46303	micrococcus
150	7	1.0	156	1	Y805_AQUAE	066989	aquifex aeo	223	7	1.0	286	1	GD40_WHEAT	P22737	synecococc
151	7	1.0	156	1	RS7_MICLU	P09898	micrococcus	224	7	1.0	286	1	CHIB_TOBAC	P02863	trititum ae
152	7	1.0	157	1	RS7_CHLNU	09pjv5	chlamydia m	225	7	1.0	294	1	GDAB_WHEAT	P29601	nicotiana t
153	7	1.0	157	1	RS7_CHLNU	092801	chlamydia p	226	7	1.0	296	1	SAPC_ECOLI	P04726	trititum ae
154	7	1.0	157	1	RS7_CHLNU	P29765	chlamydia t	227	7	1.0	296	1	SAPC_SALTU	O47624	escherichia
155	7	1.0	161	1	UCN3_HUMAN	096963	homo sapien	228	7	1.0	296	1	THRM_RAT	P36669	salmonella
156	7	1.0	162	1	CRF_XENLA	P49188	xenopus lae	229	7	1.0	297	1	GD44_WHEAT	P97532	rattus norv
157	7	1.0	163	1	YCB1_BACUN	P30906	bacteroides	230	7	1.0	301	1	HXDD_CHICK	P04724	trititum ae
158	7	1.0	164	1	NU6M_STRPU	P18390	escherichia	231	7	1.0	303	1	GDFF_MOUSE	P24344	gallus galli
159	7	1.0	165	1	NU6M_STRPU	P15553	strongyloce	232	7	1.0	303	1	GDFF_MOUSE	O92016	rattus muscucu
160	7	1.0	167	1	NU6M_MXGCL	P08447	myxine glut	233	7	1.0	304	1	CBPZ_SIMVI	P42788	simulium vi
161	7	1.0	171	1	ATPF_SYNP6	P08447	synecococc	234	7	1.0	304	1	PEX2_CRIGR	O06438	cricetulus
162	7	1.0	176	1	TR23_MOUSE	O9e163	synecococc	235	7	1.0	305	1	PEX2_HUMAN	P28328	homo sapien
163	7	1.0	178	1	RR7_ASTLO	P14760	astasia lon	236	7	1.0	305	1	PEX2_MOUSE	P55098	mus muscucu
164	7	1.0	182	1	FTN_DROME	P35554	drosophila	237	7	1.0	305	1	YCA4_LISIN	O925x2	listeria in



C 238	7	1.0	307	1	GDA9_WHEAT	P18573	triticum ae	311	7	1.0	412	1	IF42_ARATH	P41377	arabidopsis
C 239	7	1.0	313	1	GDAT_WHEAT	P04727	triticum ae	312	7	1.0	412	1	IF42_ARATH	O09671	schizosacch
C 240	7	1.0	314	1	MIAA_CHLMT	P04727	chlamydia t	313	7	1.0	413	1	IF42_TOBAC	O40468	nicotiana t
C 241	7	1.0	314	1	MIAA_CHLTR	O84771	chlamydia t	314	7	1.0	414	1	TY41_HUMAN	P25490	homo sapien
C 242	7	1.0	314	1	NMI_MOUSE	O35309	mus musculu	315	7	1.0	414	1	YANO_SCHPO	O10075	schizosacch
C 243	7	1.0	316	1	LORI_HUMAN	P23490	homo sapien	316	7	1.0	416	1	SO_DROME	O27350	drosophila
C 244	7	1.0	319	1	GDAS_WHEAT	O041725	triticum ae	317	7	1.0	420	1	PAR1_XENLA	P47749	xenopus lae
C 245	7	1.0	324	1	YI70_ARCFU	O28112	archaeoglob	318	7	1.0	423	1	SYH_PASMU	P57988	pasteurella
C 246	7	1.0	325	1	TRBB_RHISN	P55395	rhicobium s	319	7	1.0	427	1	YDF1_ECOLI	P77282	escherichia
C 247	7	1.0	325	1	YGS4_ARCFU	O28619	archaeoglob	320	7	1.0	433	1	ENO4_MOUSE	P17182	mus musculu
C 248	7	1.0	334	1	WDR5_HUMAN	O9UGP9	homo sapien	321	7	1.0	433	1	ENO4_RAT	P04764	rattus norv
C 249	7	1.0	335	1	FLIG_THEMEA	O9WY63	thermotoga	322	7	1.0	433	1	ENO_DROME	P15007	drosophila
C 250	7	1.0	337	1	FSA_SHEEP	P31514	ovis aries	323	7	1.0	433	1	HXB3_MOUSE	P09026	mus musculu
C 251	7	1.0	337	1	TALI_MOUSE	O93092	mus musculu	324	7	1.0	433	1	NU4M_RHISA	O99825	rhidipephal
C 252	7	1.0	341	1	FSA_XENLA	P31515	xenopus lae	325	7	1.0	433	1	VPAP_HCMVA	P16790	human cytom
C 253	7	1.0	343	1	FSA_CHICK	O90844	gallus gall	326	7	1.0	436	1	CCRR_MOUSE	O08786	mus musculu
C 254	7	1.0	344	1	FSA_BOVIN	P50281	bos taurus	327	7	1.0	436	1	ENO1_ENTHI	O08786	mus musculu
C 255	7	1.0	344	1	FSA_HORSE	O62650	equus caball	328	7	1.0	436	1	ENO1_NEOFR	P51555	entamoeba h
C 256	7	1.0	344	1	FSA_PIG	P47931	mus musculu	329	7	1.0	437	1	NU4M_ALBEO	P42894	neocallimias
C 257	7	1.0	344	1	FSA_PIG	P10669	sus scrofa	330	7	1.0	449	1	NU4M_CHLRE	P42894	neocallimias
C 258	7	1.0	345	1	FSA_PIG	P1674	rattus norv	331	7	1.0	445	1	ENO1_HEVBR	P20113	chlamydomon
C 259	7	1.0	345	1	YI44_CAEEL	P46500	caenorhabdi	332	7	1.0	445	1	ENO2_HEVBR	O91e10	hevea brasl
C 260	7	1.0	348	1	CHLT_EUGGR	P31205	euglena gra	333	7	1.0	445	1	ENO2_FICCO	O91e10	hevea brasl
C 261	7	1.0	348	1	N33_HUMAN	O13454	homo sapien	334	7	1.0	445	1	TEA2_MOUSE	P42896	ricinus com
C 262	7	1.0	351	1	H1B1_PASMU	O9CLM3	pasteurella	335	7	1.0	449	1	C13A_MYCTU	P48301	mus musculu
C 263	7	1.0	355	1	HKL6_LYCES	O22299	lycoperisico	336	7	1.0	450	1	Y425_MYCCE	O08447	mycobacteri
C 264	7	1.0	355	1	LAVI_PHYPO	P14725	physarum po	337	7	1.0	450	1	Y425_MYCCE	P47664	mycoplasma
C 265	7	1.0	356	1	SMRI_PODAN	O08142	podospora a	338	7	1.0	452	1	NU4M_BRATL	O47423	branchiosto
C 266	7	1.0	357	1	RUBE_STRCO	O91291	streptomyce	339	7	1.0	452	1	NU4M_BRATL	O79421	branchiosto
C 267	7	1.0	360	1	GP25_HUMAN	O00155	homo sapien	340	7	1.0	452	1	YK97_MYCLE	P54077	mycobacteri
C 268	7	1.0	361	1	ELVA_BACHD	P41363	bacillus ha	341	7	1.0	452	1	YK97_MYCTU	O10706	mycobacteri
C 269	7	1.0	361	1	WDS_DROME	O9V318	drosophila	342	7	1.0	453	1	AS16_HUMAN	O96n65	homo sapien
C 270	7	1.0	361	1	YIDS_ECO57	O8V288	escherichia	343	7	1.0	458	1	ME31_DROME	P23128	drosophila
C 271	7	1.0	366	1	IHA_MOUSE	O04997	mus musculu	344	7	1.0	458	1	YVW6_CAEEL	P34408	caenorhabdi
C 272	7	1.0	366	1	IHA_RAT	P17490	rattus norv	345	7	1.0	459	1	NU4M_PELSU	O79677	pelomedusa
C 273	7	1.0	372	1	JBH1_MOUSE	P24815	m 3 beta-hy	346	7	1.0	465	1	VP19_HSEVB	P28935	equine hecp
C 274	7	1.0	372	1	JBH1_RAT	P22071	r 3 beta-hy	347	7	1.0	466	1	GUIN5_THIEP	O01786	thermonomon
C 275	7	1.0	372	1	JBH2_RAT	P22072	r 3 beta-hy	348	7	1.0	467	1	C761_SOLME	P37121	solanum mel
C 276	7	1.0	372	1	JBH3_RAT	P27364	rattus norv	349	7	1.0	469	1	DPD2_HUMAN	P49005	homo sapien
C 277	7	1.0	372	1	JBH4_MOUSE	O61767	mus musculu	350	7	1.0	470	1	PROP_CAVPO	O64181	cavia porce
C 278	7	1.0	372	1	JBH4_RAT	O62878	r 3 beta-hy	351	7	1.0	474	1	DLDH_HALVO	O04829	halobacteri
C 279	7	1.0	373	1	SRB3_HUMAN	O9n666	homo sapien	352	7	1.0	480	1	PRTP_HUMAN	P10619	homo sapien
C 280	7	1.0	379	1	CIT2_ECOLI	P05853	escherichia	353	7	1.0	481	1	LORI_MOUSE	P18165	mus musculu
C 281	7	1.0	380	1	CATA_CHICK	P43691	gallus gall	354	7	1.0	482	1	CATA_BORPE	P48062	bordelella
C 282	7	1.0	380	1	TRPD_YEAST	O03478	saccharomyc	355	7	1.0	482	1	K6B2_HUMAN	O9ub50	h ribosomai
C 283	7	1.0	381	1	CYB_NORTY	O03478	notoryctes	356	7	1.0	482	1	TPR_FORGI	P25806	porphyromon
C 284	7	1.0	383	1	NAPR_ENTHR	P26235	enterococcu	357	7	1.0	485	1	GLR_RAT	P30082	rattus norv
C 285	7	1.0	389	1	SERI_BOMMO	P07856	bombyx mori	358	7	1.0	487	1	AMPL_MOUSE	O9c8y7	mus musculu
C 286	7	1.0	389	1	WN2B_MOUSE	O70283	mus musculu	359	7	1.0	489	1	NIFD_THIFE	P06662	thiobacillu
C 287	7	1.0	390	1	PER_DROTP	P91716	drosophila	360	7	1.0	491	1	NAR1_YEAST	P23503	saccharomyc
C 288	7	1.0	391	1	CAR2_RHINI	P43231	rhizopus ni	361	7	1.0	492	1	MM1_MOUSE	O02853	mus musculu
C 289	7	1.0	391	1	WN2B_HUMAN	O93097	homo sapien	362	7	1.0	497	1	HMS3_DROME	P18488	drosophila
C 290	7	1.0	392	1	TRE1_DROME	O9ndm2	drosophila	363	7	1.0	497	1	NU4M_ACACA	O37375	acanthamoeb
C 291	7	1.0	392	1	Y462_TREPA	O83475	treponema p	364	7	1.0	500	1	PTGI_BOVIN	O29626	bos taurus
C 292	7	1.0	394	1	ACKA_LACSK	O9x4m1	lactobacilli	365	7	1.0	501	1	C72N_ARATH	O91lmo	arabidopsis
C 293	7	1.0	396	1	CX47_HUMAN	P36383	homo sapien	366	7	1.0	501	1	CSD2_DROME	O9wnt6	drosophila
C 294	7	1.0	397	1	HKL2_MALDO	O04135	malus domes	367	7	1.0	504	1	ETS4_DROME	P29775	drosophila
C 295	7	1.0	397	1	LIRG_BOVIN	O29458	bos taurus	368	7	1.0	506	1	VE2_HPV47	P22420	human papil
C 296	7	1.0	398	1	HKL1_MALDO	O04134	malus domes	369	7	1.0	511	1	R104_PROWI	P46743	proctotecta
C 297	7	1.0	398	1	Y306_CHLMU	O9p102	chlamydia m	370	7	1.0	512	1	FUS_BOVIN	O28009	bos taurus
C 298	7	1.0	399	1	PAL1_YEAST	O12099	saccharomyc	371	7	1.0	513	1	C942_VICSA	P98188	vicia sativ
C 299	7	1.0	401	1	UXUA_BRUME	O8ycq4	brucella me	372	7	1.0	513	1	CDP2_MAIZE	P49101	zea mays (m
C 300	7	1.0	403	1	IF4A_LEIBR	O52225	leishmania	373	7	1.0	517	1	MTC1_COREO	P42828	corynebacte
C 301	7	1.0	403	1	SHBG_RAT	B08669	rattus norv	374	7	1.0	518	1	COX1_POLOS	O95511	polypterus
C 302	7	1.0	404	1	CAZ_DROME	O27254	drosophila	375	7	1.0	518	1	YH69_YEAST	P38758	saccharomyc
C 303	7	1.0	404	1	YS69_MYCTU	O33331	mycobacteri	376	7	1.0	520	1	RAG2_XENLA	O91630	xenopus lae
C 304	7	1.0	405	1	ACKA_HUMAN	P57272	buchnera ap	377	7	1.0	521	1	NPRE_BACSU	P06642	bacillus su
C 305	7	1.0	411	1	IF4N_HUMAN	P38919	homo sapien	378	7	1.0	525	1	ESR1_MICUN	P57753	micropogon
C 306	7	1.0	411	1	IHH_MOUSE	P97884	mus musculu	379	7	1.0	526	1	HYFF_ECOLI	P77437	escherichia
C 307	7	1.0	412	1	DIJ2_HUMAN	O60884	homo sapien	380	7	1.0	526	1	KIC2_BOVIN	P06394	bos taurus
C 308	7	1.0	412	1	DIJ2_MOUSE	O9qy10	mus musculu	381	7	1.0	527	1	DPE2_HUMAN	P56282	homo sapien
C 309	7	1.0	412	1	DJ42_RAT	O35824	rattus norv	382	7	1.0	527	1	NU5M_CAEEL	P24696	caenorhabdi
C 310	7	1.0	412	1	IF41_ARATH	P41376	arabidopsis	383	7	1.0	528	1	RAG2_CHICK	P25022	gallus gall

C 384	7	1.0	529	1	YB89_YEAST	P38140 saccharomyc	C 457	7	1.0	708	1	GBF_DICDI	P36417 dictyostell
C 385	7	1.0	532	1	BTB3_MOUSE	P58545 mus musculu	C 458	7	1.0	709	1	PNP_PHOLI	P41121 photorhabdu
C 386	7	1.0	532	1	CDPK_DAICA	P28582 daucus caro	C 459	7	1.0	716	1	KIF2_MOUSE	P28740 mus musculu
C 387	7	1.0	533	1	CDP2_ORISA	P53683 cryza sativ	C 460	7	1.0	718	1	MCAK_CRIGR	P70096 cricetulus
C 388	7	1.0	533	1	EGR1_MOUSE	P08046 mus musculu	C 461	7	1.0	722	1	SYGB_STNY3	O55690 eynechocyt
C 389	7	1.0	534	1	CDP1_ORISA	P53682 cryza sativ	C 462	7	1.0	725	1	MCAK_HUMAN	O99661 homo sapien
C 390	7	1.0	538	1	GLE1_YEAST	Q12315 saccharomyc	C 463	7	1.0	725	1	VR2A_BPT4	P03690 bacterioph
C 391	7	1.0	539	1	DOPE2_DROME	Q24563 drosophila	C 464	7	1.0	729	1	KAR3_YEAST	P17119 saccharomyc
C 392	7	1.0	541	1	NU57_YEAST	P48837 saccharomyc	C 465	7	1.0	730	1	KCM1_XENLA	O91635 xenopus lae
C 393	7	1.0	542	1	CDP3_ORISA	P53684 cryza sativ	C 466	7	1.0	732	1	TR16_ECOLI	O00191 escherichia
C 394	7	1.0	543	1	TCPH_HUMAN	O99832 homo sapien	C 467	7	1.0	737	1	SKN1_CANAL	P87024 candida alb
C 395	7	1.0	544	1	TCPH_MOUSE	P80313 mus musculu	C 468	7	1.0	738	1	YAS9_SCHPO	O10145 schizosacch
C 396	7	1.0	547	1	YMX1_CABEL	P34509 caenorhabd1	C 469	7	1.0	740	1	DDX1_HUMAN	O92499 homo sapien
C 397	7	1.0	549	1	DSX_DROME	P23023 drosophila	C 470	7	1.0	745	1	KATB_ARATH	P46864 arabidopsi
C 398	7	1.0	553	1	FXCI_HUMAN	Q12948 homo sapien	C 471	7	1.0	754	1	ECE1_BOVIN	P42891 bos taurus
C 399	7	1.0	553	1	FXCI_MOUSE	O61572 mus musculu	C 472	7	1.0	754	1	ECE1_CAVRO	P97739 cavia porce
C 400	7	1.0	553	1	TF3B_CANAL	P43072 candida alb	C 473	7	1.0	754	1	KATC_ARATH	P46875 arabidopsi
C 401	7	1.0	554	1	KLP3_SCHPO	O94660 schizosacch	C 474	7	1.0	759	1	C1Q2_MOUSE	O92351 mus musculu
C 402	7	1.0	557	1	G6PI_KLULA	P12341 kluyveromyc	C 475	7	1.0	762	1	ECB1_RAT	P42893 rattus norv
C 403	7	1.0	564	1	VAS5_VACCV	P21073 vaccinia vi	C 476	7	1.0	763	1	DYRA_HUMAN	O13627 homo sapien
C 404	7	1.0	564	1	VAS5_VACCV	P40954 vaccinia vi	C 477	7	1.0	763	1	DYRA_MOUSE	O61314 mus musculu
C 405	7	1.0	567	1	CH13_CANAL	P40954 candida alb	C 478	7	1.0	763	1	DYRA_RAT	O61370 rattus norv
C 406	7	1.0	572	1	SYK_CABEL	Q22039 caenorhabd1	C 479	7	1.0	767	1	ACBS_CHICK	P36196 gallus gall
C 407	7	1.0	572	1	YOG2_DEIRA	O91rm6 delnoccocus	C 480	7	1.0	770	1	ECB1_HUMAN	P42892 homo sapien
C 408	7	1.0	573	1	AMH2_HUMAN	Q16671 homo sapien	C 481	7	1.0	770	1	KLPA_EMENT	P28739 emeticella
C 409	7	1.0	577	1	CHVG_RHIME	P72292 rhizobium m	C 482	7	1.0	793	1	KATA_ARATH	O07970 arabidopsi
C 410	7	1.0	578	1	FD21_MYCTU	O50441 mycobacteri	C 483	7	1.0	793	1	KP3C_HUMAN	O14782 homo sapien
C 411	7	1.0	579	1	FD21_MYCTU	P54200 mycobacteri	C 484	7	1.0	797	1	AP32_HUMAN	O94466 homo sapien
C 412	7	1.0	580	1	YML1_ARATH	Q22752 arabidopsi	C 485	7	1.0	802	1	NAB3_YEAST	P38996 saccharomyc
C 413	7	1.0	581	1	ESR1_PAGMA	O42132 pagrus majo	C 486	7	1.0	814	1	SEFC_SALEN	P33388 salmonella
C 414	7	1.0	585	1	MPD3_HUMAN	Q13368 homo sapien	C 487	7	1.0	819	1	PPSA_PYRAB	O9vzh7 pyrococcus
C 415	7	1.0	587	1	NDCC2_RAT	P70545 rattus norv	C 488	7	1.0	821	1	PPSA_PYRHO	O9v730 pyrococcus
C 416	7	1.0	590	1	PHO4_NEUCR	P15710 neurospora	C 489	7	1.0	822	1	FER_HUMAN	P16591 homo sapien
C 417	7	1.0	590	1	SYKC_YEAST	P5180 saccharomyc	C 490	7	1.0	823	1	PARC_SMRPN	P72525 streptococ
C 418	7	1.0	593	1	KICU_HUMAN	P13645 homo sapien	C 491	7	1.0	832	1	KLPI_SCHPO	O92376 schizosacch
C 419	7	1.0	594	1	DBP9_YEAST	Q06218 saccharomyc	C 492	7	1.0	836	1	UME6_YEAST	P39001 saccharomyc
C 420	7	1.0	606	1	MM17_HUMAN	O91429 homo sapien	C 493	7	1.0	837	1	XYN2_CLOTM	P10478 clostridium
C 421	7	1.0	614	1	ZE29_MOUSE	Q07230 mus musculu	C 494	7	1.0	842	1	SYA_CAMJE	O9p105 campylobact
C 422	7	1.0	618	1	CHEA_LISIN	Q924w2 listeria in	C 495	7	1.0	844	1	DNL4_HUMAN	P49917 homo sapien
C 423	7	1.0	618	1	CHEA_LISMO	O48768 listeria mo	C 496	7	1.0	852	1	C1Q2_RAT	O88843 rattus norv
C 424	7	1.0	622	1	FACG_HUMAN	O15287 homo sapien	C 497	7	1.0	854	1	CNRB_HUMAN	P35813 homo sapien
C 425	7	1.0	624	1	DSK1_CYLFU	Q39493 cylindroche	C 498	7	1.0	854	1	VG12_BPH2	P20345 bacterioph
C 426	7	1.0	625	1	C1KG_RAT	O63734 rattus norv	C 499	7	1.0	872	1	C1Q2_HUMAN	O43526 homo sapien
C 427	7	1.0	627	1	K2C1_MOUSE	P04104 mus musculu	C 500	7	1.0	876	1	PR02_YEAST	P20095 saccharomyc
C 428	7	1.0	635	1	VP40_HSV11	P10210 herpes simp	C 501	7	1.0	878	1	C1Q5_MOUSE	O91K45 mus musculu
C 429	7	1.0	636	1	CTR4_HUMAN	O43246 homo sapien	C 502	7	1.0	878	1	FIMD_ECOLI	P30130 escherichia
C 430	7	1.0	637	1	FTSH_BACSU	P37476 bacillus su	C 503	7	1.0	884	1	SECA_OILU	O32743 olithodisc
C 431	7	1.0	643	1	CTK2_XENLA	P79955 xenopus lae	C 504	7	1.0	887	1	SM6B_RAT	O70141 rattus norv
C 432	7	1.0	645	1	PD44_HUMAN	P13667 homo sapien	C 505	7	1.0	890	1	AT58_HUMAN	O9u979 homo sapien
C 433	7	1.0	649	1	VATI_METMA	O59659 methanosarc	C 506	7	1.0	897	1	C1Q5_HUMAN	O9n182 homo sapien
C 434	7	1.0	652	1	CD93_HUMAN	O9npv3 homo sapien	C 507	7	1.0	900	1	KDPD_CLOAB	P94608 clostridium
C 435	7	1.0	652	1	SSP1_SCHPO	P50526 schizosacch	C 508	7	1.0	907	1	GACS_PESBY	P48827 pseudomonas
C 436	7	1.0	653	1	HGFA_MOUSE	O91038 mus musculu	C 509	7	1.0	928	1	VGJB_MCMVS	P27171 murine cyto
C 437	7	1.0	655	1	HGFA_HUMAN	O04756 homo sapien	C 510	7	1.0	935	1	KINH_SYNRA	O43093 syntrophalae
C 438	7	1.0	658	1	CNT1_RABIT	O9mz12 oryctolagus	C 511	7	1.0	953	1	Y442_HUMAN	O9upv9 homo sapien
C 439	7	1.0	664	1	MS16_YEAST	P15424 saccharomyc	C 512	7	1.0	955	1	KINL_LEICH	P46865 leishmania
C 440	7	1.0	665	1	BYR4_SCHPO	O10951 schizosacch	C 513	7	1.0	958	1	AMYG_DEBOC	P22661 dactyomyces
C 441	7	1.0	671	1	HMOC_DROME	P22810 drosophila	C 514	7	1.0	966	1	PIB1_PETMA	P02574 dactyomylon
C 442	7	1.0	671	1	KRP2_RAT	O62909 rattus norv	C 515	7	1.0	970	1	T3RE_BPT1	P08764 bacterioph
C 443	7	1.0	671	1	NCPR_MUSDO	O07994 musca domes	C 516	7	1.0	971	1	YO1A_CABEL	O09281 caenorhabd1
C 444	7	1.0	671	1	PEXH_YARLI	P87200 yarrowia li	C 517	7	1.0	974	1	RH66_HUMAN	O43189 haemophilus
C 445	7	1.0	673	1	OSM3_CABEL	P46873 caenorhabd1	C 518	7	1.0	981	1	GLNE_HAEIN	P44419 mycoplasma
C 446	7	1.0	673	1	KNS2_HUMAN	O9bw19 homo sapien	C 519	7	1.0	982	1	PI15_NYCGE	O92543 mus musculu
C 447	7	1.0	675	1	HS7M_PEA	P37990 pisum sativ	C 520	7	1.0	986	1	RH66_MOUSE	O54834 mus musculu
C 448	7	1.0	679	1	KIF2_HUMAN	O00139 homo sapien	C 521	7	1.0	992	1	SNXJ_HUMAN	O92543 mus sapien
C 449	7	1.0	682	1	DY12_CHLRE	O91637 chlamydomon	C 522	7	1.0	994	1	TNP5_ECOLI	O98504 escherichia
C 450	7	1.0	682	1	KIF2_XENLA	O91637 xenopus lae	C 523	7	1.0	1004	1	PTPX_RAT	O63475 rattus norv
C 451	7	1.0	683	1	VTER_HSVSA	O01020 herpesviru	C 524	7	1.0	1004	1	SAI2_MOUSE	O9q396 mus musculu
C 452	7	1.0	684	1	PLB4_HCMVA	P17151 human cytom	C 525	7	1.0	1011	1	MYIA_DROME	O23978 drosophila
C 453	7	1.0	686	1	PLB3_YEAST	O08108 saccharomyc	C 526	7	1.0	1011	1	POIG_PMDVT	P15072 focc-and-mo
C 454	7	1.0	695	1	CIQ4_HUMAN	P56686 homo sapien	C 527	7	1.0	1013	1	M2B1_MOUSE	O09159 mus musculu
C 455	7	1.0	697	1	BYN_DROME	P55965 drosophila	C 528	7	1.0	1013	1	PMPG_CHLTR	O84879 chlamydia t
C 456	7	1.0	700	1	NCD_DROME	P20480 drosophila	C 529	7	1.0	1017	1	ACAD_ARATH	O911K7 arabidopsi

530	7	1.0	1020	1	VI20_HSVB	P28956 equine hep	c 603	6	0.9	31	1	LPL_BUCRP	Q53017 buchnera ap
531	7	1.0	1023	1	GLT_DROME	P33438 drosophila	604	6	0.8	32	1	CAL_BOVIN	P01260 bos taurus
532	7	1.0	1024	1	CCP5_HUMAN	Q96xr8 homo sapien	605	6	0.8	32	1	CAL_PIG	P01259 sus scrofa
533	7	1.0	1044	1	BUB1_SCHPO	O94751 schizosacch	606	6	0.9	32	1	MDH_NITRA	P10687 nitroschia a
534	7	1.0	1063	1	ELF1_DROME	P13002 drosophila	607	6	0.8	32	1	PRTF_SCYCA	P30259 scylitichinu
535	7	1.0	1064	1	ISKS_HUMAN	Q9ng38 homo sapien	608	6	0.8	37	1	PETG_CUSRE	P30398 cuscuta ref
536	7	1.0	1078	1	GYRB_SXYV3	P77966 synchocyst	609	6	0.9	40	1	VIT_MELGA	P56341 meleagris v
537	7	1.0	1086	1	DPOB_SCHPO	P30316 schizosacch	610	6	0.9	42	1	PSBK_CHLUV	P56348 chlorella v
538	7	1.0	1093	1	AF17_HUMAN	P55198 homo sapien	611	6	0.9	43	1	MUT1_ENTNM	P80925 enterococcu
539	7	1.0	1129	1	EG27_CAEEL	O09228 caenorhabdi	612	6	0.9	44	1	PSBK_ODOSI	P49513 odontella e
540	7	1.0	1139	1	INA1_CAEEL	Q03600 caenorhabdi	613	6	0.9	45	1	PSBK_EUGVI	Q9m55 euglena vir
541	7	1.0	1143	1	RPOB_PORPU	P33174 mus musculu	614	6	0.9	46	1	PSBK_CHLRE	P18263 chlamydomon
542	7	1.0	1146	1	CCAS_RAT	O02485 rattus norv	615	6	0.9	48	1	PSBK_EUGSA	Q9m61 euglena san
543	7	1.0	1164	1	TSC1_HUMAN	Q92574 homo sapien	616	6	0.9	49	1	R331_STRPN	O97459 streptococc
544	7	1.0	1216	1	RPOB_TROMH	Q93gfc troyermya	617	6	0.8	50	1	HSP1_HUMAN	P04553 homo sapien
545	7	1.0	1235	1	KEF4_CHICK	O90640 gallus gall	618	6	0.8	50	1	HSP1_HYLLA	P35306 hylobates l
546	7	1.0	1236	1	KEF4_XENTL	Q91784 xenopus lae	619	6	0.8	53	1	RL2A_PIG	Q29933 sus scrofa
547	7	1.0	1231	1	KEF4_MOUSE	P33174 mus musculu	620	6	0.9	54	1	B2MG_MESAU	P30432 mesocricetu
548	7	1.0	1232	1	KEF4_HUMAN	O95239 homo sapien	621	6	0.8	54	1	GALA_CANFA	P33710 canis fami1
549	7	1.0	1234	1	PIP3_MOUSE	P51432 mus musculu	622	6	0.8	56	1	HSP_COTUA	P14402 coturnix co
550	7	1.0	1235	1	DNB1_HCMVA	P7147 human cytom	623	6	0.8	57	1	HSP1_DIDMA	P35305 didelphis m
551	7	1.0	1243	1	DPOA_ORYSA	O48653 oryza sativ	624	6	0.9	57	1	LHA_CHLAV	P07503 chloroflexu
552	7	1.0	1262	1	TPP2_MOUSE	O64514 mus musculu	625	6	0.9	57	1	RK32_CHLUV	P56357 chlorella v
553	7	1.0	1267	1	Y211_HUMAN	Q92610 homo sapien	626	6	0.9	58	1	RL32_NEIMA	Q9j452 neisseria m
554	7	1.0	1282	1	DOME_DROME	O9vwe0 drosophila	627	6	0.8	58	1	RL32_NEIMB	O9j450 neisseria m
555	7	1.0	1305	1	CYA9_XENTL	P98999 xenopus lae	628	6	0.8	59	1	HSP1_PERCO	O991q5 peridarcas
556	7	1.0	1306	1	ACE_HUMAN	P12821 homo sapien	629	6	0.8	60	1	HSP1_AEPBU	O991q9 aepyrymus
557	7	1.0	1319	1	MN1_HUMAN	O10571 homo sapien	630	6	0.8	60	1	HSP1_CAEPU	P42131 caenolestes
558	7	1.0	1324	1	SKUL_BOVIN	O18964 bos taurus	631	6	0.8	60	1	HSP1_DNGCO	O991q1 dendroaspis
559	7	1.0	1375	1	GFPC_STRMU	P13470 streptococc	632	6	0.8	60	1	HSP1_MACAG	P42137 macrodipus ag
560	7	1.0	1386	1	RPOD_MAPO	O06274 marcanilia	633	6	0.8	60	1	HSP1_MACGI	P42139 macrodipus gi
561	7	1.0	1400	1	RON_HUMAN	O04912 homo sapien	634	6	0.8	60	1	HSP1_MACPA	O991q3 macropus pa
562	7	1.0	1403	1	PRO_DROME	P29617 drosophila	635	6	0.8	60	1	HSP1_ONCFR	O991q8 onychogalea
563	7	1.0	1490	1	CRK7_HUMAN	O9nyv4 homo sapien	636	6	0.8	60	1	HSP1_POTLO	O991q8 potoropus lo
564	7	1.0	1515	1	YCR1_YEAST	P39109 saccharomyc	637	6	0.9	61	1	DNB1_BPDV	P13693 buggerigar
565	7	1.0	1518	1	KKK1_YEAST	P34244 saccharomyc	638	6	0.8	61	1	HSP1_BETPE	Q9g1q0 betongia p
566	7	1.0	1558	1	YK83_YEAST	P36028 saccharomyc	639	6	0.8	61	1	HSP1_VACEU	P42138 macropus eu
567	7	1.0	1559	1	BPT1_YEAST	P14772 saccharomyc	640	6	0.8	61	1	HSP1_MACRG	P42141 macropus r
568	7	1.0	1575	1	SVJ1_HUMAN	O43426 homo sapien	641	6	0.8	61	1	HSP1_PETXA	O991q7 petrogale x
569	7	1.0	1620	1	ALK_HUMAN	O9um73 homo sapien	642	6	0.8	61	1	HSP1_THYST	O991q8 thylagale s
570	7	1.0	1621	1	SN24_HUMAN	P97793 mus musculu	643	6	0.8	61	1	HSP1_TRIVU	P42152 trichosurus
571	7	1.0	1647	1	CO3_HUMAN	P51532 homo sapien	644	6	0.8	61	1	HSP_CHICK	P15340 gallus gall
572	7	1.0	1663	1	CHD1_HUMAN	P01024 homo sapien	645	6	0.9	61	1	PSBK_LOTUA	Q9bds2 lotus japon
573	7	1.0	1709	1	CHD1_MOUSE	O14646 homo sapien	646	6	0.8	61	1	RL21_PIG	P49666 sus scrofa
574	7	1.0	1711	1	CHD2_HUMAN	P40201 mus musculu	647	6	0.9	62	1	YPPF_BACSU	P50834 bacillus su
575	7	1.0	1739	1	BIG2_HUMAN	O14647 homo sapien	648	6	0.8	63	1	HSP1_HYPMU	O991q1 hyppifyrmo
576	7	1.0	1785	1	CCAS_CYPCA	O9y635 homo sapien	649	6	0.8	63	1	MIP_BOTAS	P81077 botriopsis as
577	7	1.0	1832	1	CCAS_CYPCA	P22316 cyprinus ca	650	6	0.8	64	1	HSP1_LAGHI	O951q2 lagorcheite
578	7	1.0	1873	1	CCAS_RABIT	O13698 homo sapien	651	6	0.8	64	1	RL35_STRCO	O88059 streptomyc
579	7	1.0	1873	1	CCAS_RABIT	P07293 oryctolagus	652	6	0.8	65	1	MYHB_PIG	P81271 sus scrofa
580	7	1.0	1912	1	PTPD_HUMAN	P23468 homo sapien	653	6	0.9	65	1	PHX2_MOUSE	P15972 mus musculu
581	7	1.0	1938	1	MYS_ABOIR	P24733 aequipteten	654	6	0.8	65	1	TXM7_DENAN	P80970 dendroaspis
582	7	1.0	2038	1	FSH_DROME	P13709 drosophila	655	6	0.8	66	1	YPUJ_ECOLI	P58033 escherichia
583	7	1.0	2116	1	MY52_DICDI	P08799 dicystostell	656	6	0.8	67	1	CSPA_MICLU	O30875 micrococcu
584	7	1.0	2167	1	BEM2_YEAST	P39960 saccharomyc	657	6	0.8	68	1	HSP1_PSECU	P42145 pseudochiro
585	7	1.0	2193	1	POLG_CXA16	O65900 c genome po	658	6	0.8	68	1	PLE2_PSEAM	O99p13 pseudopleur
586	7	1.0	2209	1	Y166_HUMAN	P50748 homo sapien	659	6	0.8	70	1	L2MW_ADE40	O64858 human adeno
587	7	1.0	2232	1	DPOE_YEAST	P21951 saccharomyc	660	6	0.9	70	1	PYRH_THERT	P43891 thermus the
588	7	1.0	2332	1	POLG_FMDVA	P03308 f genome po	661	6	0.9	70	1	Y1IE_ECOLI	P32492 escherichia
589	7	1.0	2332	1	POLG_FMDV1	P03306 f genome po	662	6	0.8	71	1	DNB1_POVCO	P03086 polyomaviru
590	7	1.0	2333	1	POLG_FMDV1	P03306 f genome po	663	6	0.8	71	1	LEA2_CAVPO	O91313 cavia porce
591	7	1.0	2336	1	POLG_FMDVZ	P49303 f genome po	664	6	0.9	72	1	YDAQ_ECOLI	P76057 escherichia
592	7	1.0	2368	1	ESR1_YEAST	P38111 saccharomyc	665	6	0.8	72	1	VP8_TCV	P17461 turrip crin
593	7	1.0	2415	1	SPCA_MOUSE	P08032 mus musculu	666	6	0.9	74	1	SRP_SOYBN	O07502 glycine max
594	7	1.0	2517	1	NCR2_HUMAN	Q9y618 h nucleiar f	667	6	0.8	74	1	ULI1_HSVB	P28982 equine hep
595	7	1.0	3005	1	ZFH2_DROME	P28167 drosophila	668	6	0.8	75	1	YF45_MYCTU	O10781 mycobacteri
596	7	1.0	4367	1	DYHC_NEUCR	P45443 neurospora	669	6	0.9	76	1	CXO3_CONTE	O9x210 conus texti
597	7	1.0	4427	1	PKSL_BACSU	O05470 bacillus su	670	6	0.8	77	1	DIR1_ECOLI	P21320 escherichia
598	7	1.0	4660	1	LRP2_RAT	P98158 rattus norv	671	6	0.9	77	1	ENV_SMSGV	P03384 simian sarc
599	7	1.0	5035	1	RYR1_PIG	P16960 sus scrofa	672	6	0.8	77	1	LEA2_BOVIN	O951c3 bos taurus
600	7	1.0	5037	1	RYR1_RABIT	P17176 oryctolagus	673	6	0.8	77	1	PR22_SEPOF	P80002 sepi offic
601	7	1.0	5038	1	RYR1_HUMAN	P21817 homo sapien	674	6	0.8	77	1	YF78_ARCFU	O28694 archaeoglob
602	7	0.8	15	1	CXA1_CONGE	P01519 conus geogr	675	6	0.9	77	1	YF78_ARCFU	O28694 archaeoglob

675	6	0.8	78	1	PRT1_SEPOF	P80001 sepi offic	C 749	6	0.9	100	1	RL23_ACTAC	P55839 actinobacti
677	6	0.8	78	1	VE5_HPV41	P27554 human papil	C 750	6	0.9	100	1	TX0A_ATRIL	Q95b1v8 atrax sp. 1
678	6	0.8	79	1	CSMA_CHL4U	P09928 chloroflexu	C 751	6	0.9	100	1	TX0B_ATRIL	Q95b1v7 atrax sp. 1
679	6	0.8	79	1	EF6_BACST	P23081 bacillus st	C 752	6	0.8	100	1	YGF6_HAEIN	P44062 haemophilus
680	6	0.9	79	1	OAC3_SALTI	O8285 salmonella	C 753	6	0.8	100	1	YMT5_CLOTT	Q05500 clostridium
681	6	0.9	79	1	OAC3_SALTY	P58651 salmonella	C 754	6	0.9	101	1	GRO_CRIGR	P09340 cricetus
682	6	0.8	79	1	PRU3_HUMAN	P02814 homo sapien	C 755	6	0.8	101	1	HIL6_ENSMI	P27204 ensis minor
683	6	0.9	80	1	OAG1_SALTI	O8284 salmonella	C 756	6	0.9	101	1	RS16_UREPA	Q95p61 ureaplasma
684	6	0.9	80	1	YDMC_SCHPO	O13915 schizosacch	C 757	6	0.8	101	1	VPF_STIMI	P05958 simian immu
685	6	0.9	80	1	YDMC_SCHPO	O13915 schizosacch	C 758	6	0.8	102	1	CYTI_ORYSA	P09229 oryza sativ
686	6	0.9	80	1	YHVS_WOLSU	P1877 wolfinella s	C 759	6	0.8	102	1	YB46_YEAST	P38171 saccharomyC
687	6	0.9	81	1	LUXC_VIBFI	P12748 vibrio flic	C 760	6	0.9	103	1	ATPN_HUMAN	O75964 homo sapien
688	6	0.9	81	1	OAG2_SALTY	P58650 salmonella	C 761	6	0.9	103	1	MGP_FAT	P08494 rattus norv
689	6	0.8	81	1	POC3_STYRU	P58171 syringa vul	C 762	6	0.9	103	1	SM52_RANRI	P87385 rana ridibu
690	6	0.8	81	1	POC3_BACSU	P54502 bacillus su	C 763	6	0.9	103	1	YG03_YEAST	P53097 saccharomyC
691	6	0.8	82	1	CYTA_HELAN	O10992 hellanthes	C 764	6	0.9	104	1	GRO2_RABIT	P47854 oryctolagus
692	6	0.8	82	1	P8_HUMAN	O60356 homo sapien	C 765	6	0.9	104	1	GRO_CAVPO	O55235 cavia porce
693	6	0.8	83	1	COXG_SCHPO	O94581 schizosacch	C 766	6	0.8	104	1	HSP2_CALIA	O28337 callithrix
694	6	0.9	83	1	HEBC_MOUSE	O96q21 mus musculu	C 767	6	0.8	104	1	HSP2_RAT	P11248 rattus norv
695	6	0.8	83	1	V187_BPT3	P10302 bacterioph	C 768	6	0.9	104	1	MGP_MOUSE	P15788 mus musculu
696	6	0.9	83	1	VG03_BPM2	O64199 mycobacteri	C 769	6	0.8	104	1	VPF_HV2ST	P20884 human immu
697	6	0.9	84	1	BCEI_HUMAN	O60756 homo sapien	C 770	6	0.9	105	1	CORT_HUMAN	O00230 homo sapien
698	6	0.9	84	1	COAB_BP122	P15416 bacterioph	C 771	6	0.9	105	1	RLA2_LEIR	O44010 leishmania
699	6	0.9	84	1	HEBC_HUMAN	P81172 homo sapien	C 772	6	0.9	105	1	VNEM_FXVI	P22170 foxtail mos
700	6	0.9	84	1	HEBC_RAT	O99mh3 rattus norv	C 773	6	0.8	105	1	VPF_HV2SB	P12455 human immu
701	6	0.8	84	1	POC3_OLEEU	O81092 olea europ	C 774	6	0.8	105	1	YB59_YEAST	P38311 saccharomyC
702	6	0.8	84	1	RL24_SULAC	O05633 sulfolob	C 775	6	0.8	105	1	YH06_METTH	O27741 methanobact
703	6	0.8	84	1	YENX_BPPB2	P14739 bacterioph	C 776	6	0.8	106	1	YCIH_HAEIN	P45116 haemophilus
704	6	0.8	84	1	YENX_BPPB2	P12049 bacillus su	C 777	6	0.8	107	1	CY72_ORYSA	P20907 oryza sativ
705	6	0.9	84	1	YF7A_PSEAE	P58040 pseudomon	C 778	6	0.8	107	1	ELBA_ECOLI	P75987 escherichia
706	6	0.9	85	1	FCBE_BOVIN	O9bdt7 bos taurus	C 779	6	0.9	107	1	GRO_HUMAN	P09341 homo sapien
707	6	0.9	85	1	HIS3_SULSO	O33778 sulfolob	C 780	6	0.8	107	1	HSP2_MOUSE	P07978 mus musculu
708	6	0.9	85	1	MTRG_METTH	O27225 methanobact	C 781	6	0.9	107	1	M12A_HUMAN	P19875 homo sapien
709	6	0.9	85	1	MTRG_METTH	O50774 methanobact	C 782	6	0.9	107	1	M12B_HUMAN	P19876 homo sapien
710	6	0.9	86	1	FCBE_CAVPO	O07249 cavia porce	C 783	6	0.9	107	1	PRK2_RAT	O8413 rattus norv
711	6	0.9	86	1	FCBE_HUMAN	P30273 homo sapien	C 784	6	0.9	107	1	Y240_RICCR	O9zdt5 rickettsia
712	6	0.9	86	1	FCBE_PIG	O9x626 sus scrofa	C 785	6	0.9	107	1	Y324_RICCN	O921u6 rickettsia
713	6	0.8	86	1	PAK1_SYNY3	P72712 synchocyst	C 786	6	0.9	108	1	Y454_MOUSE	P18419 mus musculu
714	6	0.9	86	1	RL23_METJA	P54016 methanococc	C 787	6	0.8	108	1	Y144_BPT4	P32179 bacterioph
715	6	0.9	87	1	IM08_YEAST	P57744 saccharomyC	C 788	6	0.8	108	1	Y187_BUCOI	P57284 buchnera ap
716	6	0.9	90	1	RK16_OENAM	P42355 oenothera a	C 789	6	0.8	108	1	YCGI_ECOLI	P76003 escherichia
717	6	0.9	90	1	VG09_BPPF1	P25135 bacterioph	C 790	6	0.8	108	1	YNI1_CAEEL	P91337 caenorhabdi
718	6	0.9	91	1	DBH_STRTR	P96045 streptococc	C 791	6	0.9	109	1	ZM33_MAI2E	O81306 zea mays (m
719	6	0.9	91	1	V179_FOPW	O9j554 fowlpox vir	C 792	6	0.9	110	1	CU15_HUMAN	P58622 homo sapien
720	6	0.8	92	1	Y432_METJA	O57874 methanococc	C 793	6	0.8	110	1	YE42_THENA	O9x1f5 thermocoga
721	6	0.9	92	1	YE21_HAEIN	P44192 haemophilus	C 794	6	0.8	111	1	YR2C_BACSU	O34527 bacillus su
722	6	0.8	93	1	RT19_MARPO	P26874 marchantia	C 795	6	0.9	112	1	CORT_RAT	O62949 rattus norv
723	6	0.9	93	1	UGR1_HUMAN	O96611 homo sapien	C 796	6	0.9	112	1	OL7F_MOUSE	O60887 mus musculu
724	6	0.9	93	1	Y435_METJA	O57877 methanococc	C 797	6	0.8	112	1	SNS4_RAT	P02783 rattus norv
725	6	0.9	94	1	Y1F8_YEAST	P40521 saccharomyC	C 798	6	0.9	112	1	SVS4_RAT	P02783 rattus norv
726	6	0.9	95	1	YCNE_BACSU	P94425 bacillus su	C 799	6	0.9	112	1	TFP1_CAVPO	P97273 cavia porce
727	6	0.8	95	1	YPS8_RHILP	P10497 rhizobium t	C 800	6	0.9	113	1	YBVO_YEAST	P31659 saccharomyC
728	6	0.9	96	1	NU6W_ALBTU	O08084 albiantia t	C 801	6	0.9	113	1	KVZD_HUMAN	P08157 homo sapien
729	6	0.9	97	1	7E_DICDI	P15649 dicystosteli	C 802	6	0.8	114	1	IR03_HCWA	P17142 human cytom
730	6	0.8	97	1	CY7T_VTGUN	O06445 vigina ungu	C 803	6	0.9	114	1	NPEF_MOUSE	O9wva8 mus musculu
731	6	0.8	97	1	DEF4_HUMAN	P12838 homo sapien	C 804	6	0.9	114	1	NPEF_RAT	O9wva8 rattus norv
732	6	0.9	97	1	IM8A_RAT	O9wva1 rattus norv	C 805	6	0.9	114	1	SZ05_HUMAN	P42830 homo sapien
733	6	0.9	97	1	VE7_HPV23	P50781 human papil	C 806	6	0.9	114	1	S206_HUMAN	P80167 homo sapien
734	6	0.8	97	1	VEP_SIAWK	P05957 simian immu	C 807	6	0.9	115	1	ATKA_BACST	P44456 b potassium
735	6	0.8	97	1	YED6_HAEIN	P23554 caldocellum	C 808	6	0.8	115	1	HIS3_MYCTU	O947c3 mycobacteri
736	6	0.8	97	1	YX1B_CALSA	P23554 caldocellum	C 809	6	0.8	115	1	HIS3_MYCTU	O53909 mycobacteri
737	6	0.8	98	1	KRFT_LARNO	P02451 iarus novae	C 810	6	0.9	115	1	RA2A_ARATH	Q961f7 arabidopsia
738	6	0.9	98	1	NU1M_LUMTE	O34948 lumbricus t	C 811	6	0.9	115	1	RA2B_ARATH	P51407 arabidopsia
739	6	0.9	98	1	SNC5_HUMAN	O75971 homo sapien	C 812	6	0.9	115	1	SM51_PRONN	O9w7f0 picropleurus
740	6	0.9	98	1	V55_BPT3	P20319 bacterioph	C 813	6	0.9	115	1	Y440_ARCFU	O29609 archaeoglob
741	6	0.9	99	1	B2MG_RABIT	P01885 oryctolagus	C 814	6	0.9	115	1	YNA3_YEAST	P53827 saccharomyC
742	6	0.9	99	1	SR19_PYRAB	O9v1j9 pyrococcus	C 815	6	0.8	115	1	YSCI_YEREN	Q001250 yerisina en
743	6	0.9	99	1	SR19_PYRAB	O9v1j9 pyrococcus	C 816	6	0.8	115	1	YSCI_YERPE	Q000933 yerisina pe
744	6	0.9	99	1	SZ14_HUMAN	O95715 homo sapien	C 817	6	0.8	116	1	GLB_PARCA	P15160 paramecium
745	6	0.9	99	1	SZ14_MOUSE	O9wvcs mus musculu	C 818	6	0.8	116	1	NU3M_ASTPE	P11591 asterixina pe
746	6	0.9	99	1	YHCO_BACSU	P54599 bacillus su	C 819	6	0.9	116	1	NU3M_STRCA	O79102 struthio caa
747	6	0.9	100	1	BOP1_BOVIN	P83107 bos taurus	C 820	6	0.8	116	1	Y601_HAEIN	P44084 haemophilus
748	6	0.9	100	1	W12B_RAT	Q10747 rattus norv	C 821	6	0.9	116	1	YB01_YEAST	P38233 saccharomyC

C 822	6	0.9	117	1	GP49_BPSPI	048403 bacterioph	895	6	0.8	132	1	ATPE_CVAPA	P48083 cyanophora
C 823	6	0.9	117	1	KV2E_HUMAN	P06309 homo sapien	C 896	6	0.9	132	1	Y125_AQUAE	O66525 aquifex ae
C 824	6	0.9	117	1	RNPA_LACIA	O9c379 lactococcus	C 897	6	0.9	132	1	YVGA_VACCC	P20565 vaccinia v1
C 825	6	0.9	117	1	UREF_BACPA	O45345 bacillus pa	C 898	6	0.8	133	1	FOJB_MYCTU	O06275 mycobacteri
C 826	6	0.9	117	1	MAP_CAMDR	P09837 camelus dro	C 899	6	0.9	133	1	NIKR_SUTSO	O9ux49 eulfolobus
C 827	6	0.8	118	1	RR15_ASTLO	P34772 astasia lon	C 900	6	0.9	133	1	SECR_MOUSE	O08535 mus musculu
C 828	6	0.9	118	1	YAI9_SCHPO	O09896 schizosacch	C 901	6	0.9	133	1	YJ23_YEAST	P47094 saccharomyc
C 829	6	0.9	119	1	HIS3_SUITO	O24413 zea mays (m	C 902	6	0.9	134	1	SECR_RAT	P11384 rat mus norv
C 830	6	0.9	119	1	RLA3_MAIZE	O97096 sulfolobus	C 903	6	0.8	134	1	STPA_ECOLI	P30017 escherichia
C 831	6	0.8	119	1	RNL4_BOVIN	P15467 bos taurus	C 904	6	0.8	134	1	YFFE_SCHPO	O14069 schizosacch
C 832	6	0.9	119	1	RNPA_BACHD	O9rc64 bacillus ha	C 905	6	0.9	134	1	YFID_BACSU	P54720 bacillus su
C 833	6	0.9	119	1	SZ07_PIG	P43030 sus scrofa	C 906	6	0.9	135	1	ANFC_SQUAT	P41319 squibus aca
C 834	6	0.8	119	1	WNT4_PLEJO	P28105 plethodon j	C 907	6	0.9	135	1	RK16_ARATH	P56793 arabidopsis
C 835	6	0.9	119	1	YJ98_YEAST	P47162 saccharomyc	C 908	6	0.9	135	1	YGDB_ECOLI	P08370 escherichia
C 836	6	0.9	120	1	CHH1_PENUP	O15960 penaeus jap	C 909	6	0.8	136	1	GLB3_CHITP	P22431 chironomus
C 837	6	0.9	120	1	CHH2_PENUP	O9u5d2 penaeus jap	C 910	6	0.9	136	1	RK16_MAIZE	P08528 zea mays (m
C 838	6	0.9	120	1	GLHA_CALJA	P51493 c glycoprot	C 911	6	0.9	136	1	RK16_ORYZA	P12138 oryza sativ
C 839	6	0.9	120	1	SY27_MOUSE	O9z1x0 mus musculu	C 912	6	0.8	137	1	PER1_ASCSU	P49671 ascaris suu
C 840	6	0.9	120	1	TR17_HUMAN	O15651 homo sapien	C 913	6	0.9	137	1	RS9_SYNY3	P73393 synchocyst
C 841	6	0.8	120	1	YANI_RHISN	P55581 rhizobium s	C 914	6	0.8	138	1	GC5H_PYRAB	O9v0g1 pyrococcus
C 842	6	0.9	121	1	AMEL_ORNAN	O97646 ornithosyn	C 915	6	0.9	138	1	LSHB_MACRU	O46483 macropus ru
C 843	6	0.9	121	1	AMEL_TACAC	O97647 tachylosyn	C 916	6	0.8	138	1	RS8_THETH	P24319 thermus ru
C 844	6	0.8	121	1	RL20_CHLIPN	O9z6f7 chlamydia p	C 917	6	0.8	138	1	TERD_ECOLI	P28816 escherichia
C 845	6	0.9	121	1	SECR_HUMAN	P09683 homo sapien	C 918	6	0.9	139	1	UGR1_MOUSE	O9z0h1 mus musculu
C 846	6	0.8	121	1	YG18_YEAST	P53151 saccharomyc	C 919	6	0.9	139	1	YN06_YEAST	P53442 saccharomyc
C 847	6	0.8	122	1	RK14_PINTH	P41633 pinus thunb	C 920	6	0.8	139	1	Y072_RHIME	P42879 rhizobium m
C 848	6	0.9	123	1	C59A_MOUSE	O55186 mus musculu	C 921	6	0.9	140	1	ANF_CHICK	P18908 gallus gall
C 849	6	0.9	123	1	RNPA_STRPN	O97n15 streptococc	C 922	6	0.9	140	1	MOPE_BRMCA	P34199 erwina car
C 850	6	0.9	123	1	UR2_MOUSE	O9qzq3 rat musculu	C 923	6	0.8	140	1	VE6_HPV04	O07854 human papil
C 851	6	0.9	123	1	UR2_RAT	O9qzq4 rattus norv	C 924	6	0.8	140	1	VE6_HPV65	O07856 human papil
C 852	6	0.9	123	1	WNT7_STRPU	P28098 strongyloce	C 925	6	0.8	140	1	Y454_MYCPN	P75127 mycoplasma
C 853	6	0.9	123	1	YPPE_BACSU	P50803 bacillus su	C 926	6	0.9	140	1	Y400_RHIME	O52667 rhizobium m
C 854	6	0.8	124	1	R22B_ARATH	O9m9w1 arabidopsis	C 927	6	0.9	140	1	YGK8_YEAST	P53139 saccharomyc
C 855	6	0.8	124	1	R22C_ARATH	O9f6e8 arabidopsis	C 928	6	0.8	140	1	YJ14_YEAST	P47111 saccharomyc
C 856	6	0.9	124	1	RNP_GIRCA	P00662 giraffa cam	C 929	6	0.8	141	1	DUT_CHVPI	O41033 paramecium
C 857	6	0.8	124	1	WNT3_EVATR	P28090 evaeterias	C 930	6	0.8	141	1	HBA_EUDSC	P19789 eudynams s
C 858	6	0.9	125	1	ALR_MOUSE	P56213 mus musculu	C 931	6	0.9	141	1	LSHB_HUMAN	P01229 homo sapien
C 859	6	0.9	125	1	ALR_RAT	O63042 rattus norv	C 932	6	0.9	141	1	LSHB_TRIUV	O46482 trihoosus
C 860	6	0.9	125	1	RS13_RICCN	O92g48 rickettsia	C 933	6	0.9	141	1	NIKR_METUA	O57965 methanococc
C 861	6	0.8	125	1	RS13_RICPR	O9zcg7 rickettsia	C 934	6	0.9	142	1	CU07_MOUSE	P58500 mus musculu
C 862	6	0.9	125	1	RS13_RICPR	O9zcg7 rickettsia	C 935	6	0.9	142	1	HBA_CHEKU	P80270 chelodonic
C 863	6	0.8	125	1	RS6_CAMOE	O9zab3 campylobact	C 936	6	0.9	142	1	IL3_CALJA	Q28334 callithrix
C 864	6	0.8	125	1	YAI7_CABEL	O20507 caenorhabdi	C 937	6	0.9	142	1	IL3_SAGOE	P51445 saguinus oe
C 865	6	0.8	126	1	RBFA_TREPA	O83860 treponema p	C 938	6	0.9	142	1	MXIM_SHIFL	O06083 shigella fl
C 866	6	0.9	127	1	COAT_BBPY7	P03630 bacterioph	C 939	6	0.8	142	1	VE6_HPV48	O80920 human papil
C 867	6	0.9	127	1	CRCB_CAUCR	O9a6v2 caulobacter	C 940	6	0.9	142	1	YBEC_ECOLI	P46888 escherichia
C 868	6	0.9	127	1	WAP_RABIT	P09412 oryctolagter	C 941	6	0.8	143	1	CAL_SHEEP	P01261 ovis aries
C 869	6	0.9	128	1	IF5A_ARCFU	O29612 archaeoglob	C 942	6	0.8	143	1	IL3_MACMU	P25140 macaca mula
C 870	6	0.9	128	1	PRK2_MOUSE	O9qxu7 mus musculu	C 943	6	0.8	143	1	RK16_MARPO	P06383 marchantia
C 871	6	0.9	128	1	YS98_MYCTU	O10819 mycobacteri	C 944	6	0.8	143	1	Y363_AQUAE	O66888 aquifex ae
C 872	6	0.8	129	1	AAK2_PIG	O28948 sus scrofa	C 945	6	0.8	143	1	YTYG_BACSU	P40779 bacillus su
C 873	6	0.9	129	1	C59E_MOUSE	P58019 mus musculu	C 946	6	0.9	144	1	CSF2_HUMAN	P04141 homo sapien
C 874	6	0.8	129	1	LYC2_CAPHI	P37714 capra hircu	C 947	6	0.9	144	1	IL3_BOVIN	P49875 bos taurus
C 875	6	0.9	129	1	LYC_SHEEP	P17607 ovis aries	C 948	6	0.8	144	1	RIB1_POHLE	P01994 photobacter
C 876	6	0.9	129	1	PRK2_HUMAN	O9hc23 homo sapien	C 949	6	0.8	144	1	SOD1_HALME	O08461 halobacteri
C 877	6	0.9	129	1	TKN1_HUMAN	P20366 homo sapien	C 950	6	0.8	144	1	YUEK_BUCAP	O44634 buchnera ap
C 878	6	0.9	129	1	VL51_CAEEL	P34386 caenorhabdi	C 951	6	0.8	145	1	ANGR_MOUSE	O64438 mus musculu
C 879	6	0.8	130	1	CW36_YEAST	P25803 saccharomyc	C 952	6	0.9	145	1	COPY_ENTHR	O47839 enterococcu
C 880	6	0.8	130	1	GTH1_CAVAU	O98848 carassius a	C 953	6	0.9	145	1	PA2W_CAVPO	P47711 cavia porce
C 881	6	0.9	130	1	OREX_CANFA	O9g1f6 canis famli	C 954	6	0.9	145	1	RL9_UREPA	O9p9t9 ureoplasm
C 882	6	0.9	130	1	OREX_MOUSE	O55241 mus musculu	C 955	6	0.8	145	1	SSRP_MYCPU	O986q9 mycoplasma
C 883	6	0.9	130	1	OREX_RAT	O55232 rattus norv	C 956	6	0.9	145	1	YA01_RHIME	O52968 rhizobium m
C 884	6	0.9	130	1	TKN1_MESAU	O60541 mesocricetu	C 957	6	0.8	145	1	YG13_SYNY3	P73361 synchocyst
C 885	6	0.9	130	1	TKN1_MOUSE	P41539 mus musculu	C 958	6	0.9	146	1	FUR2_MYCFO	O69451 mycobacteri
C 886	6	0.9	130	1	TKN1_RAT	P06767 rattus norv	C 959	6	0.9	146	1	IL3_SHEEP	O06435 ovis aries
C 887	6	0.8	130	1	YH45_YEAST	P38752 saccharomyc	C 960	6	0.8	146	1	XKDJ_BACSU	P54330 bacillus su
C 888	6	0.9	131	1	OREX_HUMAN	O43612 homo sapien	C 961	6	0.8	146	1	YD82_METUA	O58777 methanococc
C 889	6	0.9	131	1	OREX_PIG	O77668 sus scrofa	C 962	6	0.8	146	1	YG5G_YEAST	P53321 saccharomyc
C 890	6	0.9	131	1	RS11_HELPJ	O9zjtc3 helicobacte	C 963	6	0.9	147	1	CSF9_HUMAN	O9h4g1 homo sapien
C 891	6	0.9	131	1	SECR_HELPJ	P56018 helicobacte	C 964	6	0.9	147	1	FUR_MYCTU	O07724 mycobacteri
C 892	6	0.9	131	1	SECR_PIG	P01279 sus scrofa	C 965	6	0.9	147	1	LYC_COTUA	P00701 coturnix co
C 893	6	0.9	131	1	Y209_PYRHO	O57948 pyrococcus	C 966	6	0.9	147	1	MCP_MEDSA	P42067 medicago sa
C 894	6	0.8	131	1	YAOG_SCHPO	Q10095 schizosacch	C 967	6	0.8	147	1	PL1A_AGBXL	P82142 agkistrodon

968	0.8	1.7	1	PL1A	TR1FL	P2175	timeresunru
969	0.8	1.7	1	PL1E	TR1FL	P2175e	timeresunru
970	0.8	1.7	1	RL2A	HUMAN	P46776	mus sapien
971	0.8	1.7	1	RL2A	MOUSE	P1415	mus musculu
972	0.8	1.7	1	RL2A	RA1	P1845	rattus norv
973	0.8	1.7	1	RL2A	XENLA	P34930	xenopus lae
974	0.8	1.7	1	RNL4	HUMAN	P47096	mus sapien
975	0.8	1.7	1	RNL4	PIG	P15468	sus scrofa
976	0.8	1.7	1	YZ3B	MENJA	P60257	methanococ
977	0.9	1.8	1	GVPU	ANSP	P8YU1	anabaena sp
978	0.9	1.8	1	NDKE	FLABI	P47920	flaveria bi
979	0.8	1.8	1	NIKR	HELPJ	P9ZJ37	helicobacte
980	0.8	1.8	1	NIKR	HELPJ	P05589	helicobacte
981	0.8	1.8	1	RL2A	DICD1	P48160	dicrostellei
982	0.8	1.8	1	RL2A	TEMPO	P48161	euphorbia cr
983	0.8	1.8	1	RL2A	TEMPO	P27021	tendbrio mo
984	0.9	1.8	1	SY02	MOUSE	P10148	mus musculu
985	0.9	1.8	1	SY02	RA1	P10148	mus musculu
986	0.9	1.8	1	YC4Q	PORPU	P11494	rattus norv
987	0.9	1.9	1	LRP1	BACSU	P51350	porphyra pu
988	0.8	1.9	1	LRP1	BACSU	P96653	bacillus su
989	0.8	1.9	1	RBPA	CAVCR	P07661	oryza sativ
990	0.8	1.9	1	RL2A	DROME	P09630	caulobacter
991	0.8	1.9	1	RL9	HELPU	P41092	drosophila
992	0.8	1.9	1	RL9	HELPU	P41092	drosophila
993	0.9	1.9	1	YFEZ	ECOLI	P9Z1W3	helicobacte
994	0.9	1.9	1	AVR1	CHICK	P56035	helicobacte
995	0.9	1.9	1	AVR2	CHICK	P76538	eschericchia
996	0.9	1.9	1	AVR3	CHICK	O13153	gallus galli
997	0.9	1.9	1	AVR4	CHICK	P56732	gallus galli
998	0.9	1.9	1	AVR4	CHICK	P56733	gallus galli
999	0.9	1.9	1	AVR6	CHICK	P56734	gallus galli
1000	0.9	1.9	1	AVR7	CHICK	P56735	gallus galli
	0.9	1.9	1	AVR7	CHICK	P56736	gallus galli
	0.9	1.9	1	RNKC	CERAE	O46534	carotiphtho

## COMMENTS

RESULT 1	
DDX4_HUMAN	
ID_DDX4_HUMAN	STANDARD; PRT; 724 AA

AC Q9N010; Q9NBY1; Q9NT92;  
 AT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DEAD-box protein 4 (VASA homolog) .  
 GN DDX4 OR VASA.  
 OS Homo sapiens (Human) .  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NX NCBI\_TaxId=9606;  
 [1]  
 SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RP TISSUE=Testis;  
 RC MEDLINE=20402578; PubMed=10920202;  
 RX Cestrillon D.H., Quade B.J., Wang T.Y., Outgley C., Crum C.P.;  
 RT "The human VASA gene is specifically expressed in the germ cell  
 RT lineage.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9585-9590(2000) .  
 RN [2]  
 SEQUENCE FROM N.A.  
 RP Rocha D., Affara N.;  
 RT "Cloning and characterization of the human VASA gene.";   
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 SEQUENCE OF 90-724 FROM N.A.  
 RP TISSUE=Testis;  
 RC Blum H., Baerensachs S., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: May play a role in germ cell development.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- TISSUE SPECIFICITY: Expressed only in ovary and testis. Expressed  
 CC in migratory primordial germ cells in the region of the gonadal  
 CC ridge in both sexes.

```

CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC DDX4/VASA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AY004154; AAF86585.1; -
DR EMBL: AF262962; AAF72705.1; -
DR EMBL: AL137462; CAB70750.1; -
DR HSSP: Q58083; IHV8.
DR GeneW: HGNC:18700; DDX4.
DR MIM: 605281; -
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD_box.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00270; DEAD; 2.
DR Pfam: PF00271; helicase_C; 2.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00480; HELIC; 1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
DR Developmental Protein: Hydrolyase; ATP-binding; Repeat; Helicase.
FT NP_BIND 332 339 ATP (POTENTIAL).
FT SITE 446 449 DEAD BOX.
FT DOMAIN 58 234 GLY-RICH.
FT CONFLICT 556 556 I -> T (IN REF. 1).
FT CONFLICT 713 713 Q -> R (IN REF. 1).
SQ SEQUENCE 724 AA; 79307 MW; 6D313DD98B177067 CRC64;

Alignment Scores:
Pred. No.: 0 Length: 724
Score: 611.00 Matches: 711
Percent Similarity: 99.86% Conservative: 0
Best Local Similarity: 99.86% Mismatches: 1
Query Match: 84.39% Indels: 0
DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x DDX4_HUMAN (1-724)
QY 1 ATGGGGAGTGAAGATTGGGAAGCAGAAATCAACCCCTCATATGTCCTTCTATGTTCCGATA 60
Db 1 MetcilyaspgluaprtpgluaiagluileasnprohismetserSeryvalproile 20
QY 61 TTTCAGAGAGATAGTATTTCTCGAGAAATATGGACAAATTTTAAACAGACTCCAGCTTCA 120
Db 21 PhegluluyasparglyrSerglygluasnglyAspaspnheasnarythrProalaser 40
QY 121 TCATCAGAAATGGATGTATGAGCCCTTCGAAAGATCATTTTCATGAAGAAGGATTGGCC 180
Db 41 SerSerlunecaspaspaglylProSerArgraspaspisPheueclysSerglyPheala 60
QY 181 TCTGGGCGGAATTTTGGAAACAGAGATGCTGAGTGTATTAAGCGAGATATATACATCC 240
Db 61 SercllyarasnphreglyaenargaspalaglyclucysasnlysaGaapaenthSer 80
QY 241 ACAATGGCTGTTTGGAGTTGGAAAGAGTTTGGAAACAGAGGTTTTCATAACACAGAG 300
Db 81 ThrmetcllyglyPheglyValglylylYserPhegllyasnrgglyPheSerArasnserArg 100
QY 301 TTTCGAAGATGTATTAACCTCTGCTTTCTTGGAGAGAGCTGTATTAAGCTCGAAGATAT 360
Db 101 PheglunaspglyaspSerSerglylPheTrpArggluSerSerasnaspCyvgluAspsh 120
QY 361 CCAACACGGAACAGAGAGGTTTCCAAAGAGAGCGGCTATCGAGATGGAATATATTCAGAA 420
Db 121 ProthrtrrgasnargglyPheSerlYarnglycllyYtrtrrgaspglyaenasnsercllu 140
QY 421 GCTTCAGGAGCCATACAGAAAGAGGTGGAAGAGTATGTTCCGAGGTTCCGTCGAGATTT 480

```

Db 141 A1aSerGlyProTyrArgArgGlyGlyAArgGlySerPheArgGlyCysArgGlyGlyPhe 160  
 QY 481 GGCTGAGGAAGTCCAAATATGACTTAAAGCCAGAGCAATGTATGACGCGCATGGTGGC 540  
 Db 161 G1yLeuG1ySerProAsnAsnAspLeuAspProAspG1yCysMetGlnArgThrG1yGly 180  
 QY 541 CTTTGTGGTCTTGAAGACCACTATTAAAGTGCAGAGTAAAGTAACTTCTCAAGC 600  
 Db 181 LeuPheG1ySerArgArgProValLeuSerG1yThrG1yAsnG1yAspThrSerGlnSer 200  
 QY 601 AGAAGTGCAGTGGAGTGAACAGAGTGTCTTAAAGGTTAAATGAAGAAATTAACA 660  
 Db 201 ArgSerG1ySerG1ySerG1yArgG1yG1yTyrG1yG1yLeuAsnG1yGlnG1yVal11IleThr 220  
 QY 661 GGCTCGGAAGAATCTTGAAGTCAAGAGCAGAGGAGAGAGAAAGTATGATCTCA 720  
 Db 221 G1ySerG1yLysAsnSerThrLysSerG1yAlaG1yG1yG1yValG1yLysSerAspThrGln 240  
 QY 721 GGACCAAAAGTGAACATACATACCCCTCTCCACCTGAGAGTGAAGATCTCCATCTTGGCA 780  
 Db 241 G1yProLysValThrTyrIleProProProProG1yAspG1yLysSerIlePheAla 260  
 QY 781 CATTAACAGACAGCGCAATACTTGCACAAATACGACACTATTTGTGGAAGTGTCTGA 840  
 Db 261 H1sTyrG1yThrG1yLysAsnPheAspLysTyrAspThrIleLeuValG1yVal1SerGly 280  
 QY 841 CATGATGACACACAGCAATTCGACTTTGAAGAGCTAATCTCTGTCAGACACTGAAT 900  
 Db 281 H1sAspAlaProProAla1IleLeuThrPheG1yGlnAlaAsnLysG1yThrIleLeuAsn 300  
 QY 901 AACCAATTCCTAAAGCTGTATTAAGTAACTTACTCTGTGCAAAATACAGTATTTCT 960  
 Db 301 AsnAsn1IleAlaLysAlaG1yTyrThrLysLeuThrProValG1yLysSerIlePro 320  
 QY 961 ATCATCTTGCAGAGCAGAGTGTGATGCTTGTCTCAAAAGAGGTGGGAAGACTGGC 1020  
 Db 321 I1e1IleLeuAlaG1yArgAspLeuMetAlaCysAlaG1yThrG1ySerG1yLysThrAla 340  
 QY 1021 GCTTCTCTCATCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 Db 341 A1aPheLeuLeuPro1IleLeuAlaH1sMetCysH1sAsp1y1IleThrAlaSerAspPhe 360  
 QY 1081 AAAGAGTTGAGAGACAGAGTATATATGATGACACCACTGAGAAATGTCACAG 1140  
 Db 361 LysG1yLeuG1yGlnIleuProG1yCys1Ile1IleValAlaProThrArgLeuLeuValAsnGln 380  
 QY 1141 ATTATTTGGAGAGCCGAAATTTCTTTGGAGCTTGTGTAAAGCTGTTGTTATATAT 1200  
 Db 381 I1eTyrLeuG1yAlaArgLysPheSerPheG1yThrCysValArgAlaVal1IleTyr 400  
 QY 1201 GGGGGAACCCAGCTGGGACATTCATTCAGCAATGTACAAAGCTGTAATATATATGT 1260  
 Db 401 G1yG1yThrG1yLeuG1yLysSer1IleArgG1y1IleAlaG1yCysAsn1IleLeuCys 420  
 QY 1261 GCTACTCTGGAAGACTGATGATATCATGAGCAAAAGAAAGATGGTCTCAAGAGATC 1320  
 Db 421 AlaThrProG1yArgLeuMetAsp1Ile1IeG1yLysG1yLys1IleG1yLeuLysG1y1Ile 440  
 QY 1321 AATATCTAGTTTGGATGAAGCTGATCGATGTGATATGGGTTTGGTCCAGAAATG 1380  
 Db 441 LysTyrLeuValLeuAsnG1yAlaAspArgMetLeuAsnPheG1yPheG1yProG1yLysMet 460  
 QY 1381 AAGAAGTAAATTTCTGGCCAGGAATGCCATCAAGGAAGAGGCGCAACCCCTTAATGTC 1440  
 Db 461 LysLysLeuLysSerCysProG1yMetProSerLysG1yGlnIleArgIleThrLeuMetPhe 480  
 QY 1441 AGTGAACCTTTTCCAGAGAAATCAAGGTTGGCTGCAAGGTTTAAAGTCAAAATTAAT 1500  
 Db 481 SerAlaThrPheProG1yGlnIleG1yArgLeuAla1IleAlaIleuPheLysSerAsnTyr 500  
 QY 1501 CTGTTTGTCTGCTTGGAGCAAGTGGGTGGAGCATGTAGAGATGTTCAAGCAGACCGTTCTC 1560  
 Db 501 LeuPheValAlaValG1yGlnValG1yG1yAlaCysArgAspValG1yGlnIleThrValLeu 520

QY 1561 CAAGTGGCCAGTTCTCAAAAAAGAGAAAGCTCGTTGAATTTCTGCGAAAACATAGGGGAT 1620  
 Db 521 GlnValG1yGlnPheSerLysArgLysLysLeuValG1yLysLeuArgAsn1IeG1yAsp 540  
 QY 1621 GAAAGCAATAGGCTCTTTGTTGAACATAAGAAAAAGACAGATTTTACTGCACATTTTCTT 1680  
 Db 541 G1yArgThrMetValPheValG1yThrLysLysLysAlaAspPhe1IleAlaThrPheLeu 560  
 QY 1681 TGTCAAGAAAAATATCACTCAACTATCCATGCGATGCGGAAGAGAGAGCGGGAG 1740  
 Db 561 CysG1yGlnLysLysSerThrThrSer1IleH1sG1yAspArgG1yGlnIleArgG1yVal 580  
 QY 1741 CAAGCTTGAAGATTTGCTGTTGGAAGTCCCGCAGTCTTGTGCTACTTCACTAGTAGCT 1800  
 Db 581 GlnAlaLeuG1yAspPheArgPheG1yLysCysProValLeuValAlaThrSerValAla 600  
 QY 1801 GCCAGAGGCTGGATATTTGAAAAATGTCAACATGTTATCAATTTGATCTTCTTACC 1860  
 Db 601 A1aArgG1yLeuAsp1IleG1yLeuValG1yH1sVal1IleAsnPheAspLeuProSerThr 620  
 QY 1861 ATTGATGAATATGTCATGCAATTTGGCGCTACTGCTGCTGTGGGAATCTGCGCAGACA 1920  
 Db 621 IleAspG1yLysValH1sArg1IleG1yArgThrG1yArgCysG1yAsnThrG1yArgAla 640  
 QY 1921 ATTTCTTTTTCATCTTGAATCGATTAACCAATTTGACAGCGCTCTAGTAAAGTATTG 1980  
 Db 641 I1eSerPhePheAspLeuLysSerAspAsnH1sLeuAlaG1yProLeuValLysValLeu 660  
 QY 1981 ACAGATGCTCAACAGAGATGCTCTGCATGTTGGAAGAAATGCTTATGATCATCAT 2040  
 Db 661 ThrAspAlaG1yGlnG1yAspValProAlaTyrLeuG1yGln1IleAlaPheSerThrTyrIle 680  
 QY 2041 CCTGGCTTACGTGTGTATCAAGAGAAACGTTTGTGATCACTGATATACCAAGAGGC 2100  
 Db 681 ProG1yPheSerG1ySerThrArgLysAsnValPheAlaSerValAspThrArgLysGly 700  
 QY 2101 AAGAGCACTTTGAACACAGACTGGGTTTCTTCTTCA 2136  
 Db 701 LysSerThrLeuAsnThrAlaG1yPheSerSerSer 712  
 Db 712 LysSerThrLeuAsnThrAlaG1yPheSerSerSer 712  
 RESULT 2  
 DDX4 MOUSE STANDARD: PRT: 702 AA.  
 ID DDX4 MOUSE STANDARD: PRT: 702 AA.  
 AC 061456; OS05X7.  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DEAD-box protein 4 (VASA homolog) (MvH).  
 GN DDX4 OR VASA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI TaxId=10090;  
 RN [1] -  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai U., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,  
 RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glaser C., King B., Kochiwa H.,  
 RA Koehl L., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,  
 RA Blake J., Boffelli P., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustlich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombert P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";   
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE OF 60-702 FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Testis;  
 RX MEDLINE=95083681; PubMed=7991615;  
 RA Fujiwara Y., Komiya T., Kawabata H., Sato M., Fujimoto H.,  
 RA Fujiwara M., Noce T.;  
 RT "Isolation of a DEAD-family protein gene that encodes a murine  
 RT homolog of Drosophila vasa and its specific expression in germ cell  
 RT lineage.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12258-12262(1994).  
 CC -1- FUNCTION: MAY play a role in germ cell development.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERINUCLEAR.  
 CC -1- TISSUE SPECIFICITY: Testis.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN SPERMATOGENIC CELLS FROM THE  
 CC -1- SPERMATOCYTE STAGE TO THE ROUND SPERMATID STAGE.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
 CC DDX4/VASA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AK014844; BAB29578.1; -;  
 DR EMBL; D14859; BAA03584.1; -;  
 DR HSSP; Q58083; IHV8.  
 DR MGD; MG1102670; Ddx4.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 2.  
 DR Pfam; PF00271; helicase\_C; 2.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC\_C; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 KM Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.  
 FT NP\_BIND 305 312  
 FT SITE 419 422  
 FT DOMAIN 58 207  
 FT CONFLICT 152 152  
 FT CONFLICT 156 160  
 FT CONFLICT 277 277  
 FT CONFLICT 291 291  
 FT CONFLICT 368 368  
 FT CONFLICT 423 423  
 FT CONFLICT 430 430  
 FT CONFLICT 433 433  
 FT CONFLICT 448 448  
 FT CONFLICT 472 472  
 FT CONFLICT 499 509  
 FT CONFLICT 540 540  
 FT CONFLICT 680 702  
 SQ SEQUENCE 702 AA; 76470 MW; SDLEACQBDP39748 CRC64;  
 FT  
 Alignment Scores:  
 Pred. No.: 4.15e-74 Length: 702  
 Score: 82.00 Matches: 82  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.33% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-714-865-15 (1-2172) x DDX4\_MOUSE (1-702)

QY 718 CAAGACCAAAAGTACCTACATACCCCTCTCCAGCTAGAGTACGATCCATCTTT 777  
 Db 213 GINGLYPIrolysalthrlytlePProPProPProGluAspLysPserIleP 232  
 QY 778 GCACATTATCAGACGACGATTAACCTTGACAAATACGACACTATTTGTGGAAGTCT 837  
 Db 233 AAlhIeTyrGlnThrIlyleAsnPhaePlySTYrAsPThrIleuValGluValSer 252  
 QY 838 GGACATGATGACACACACCAATTCGACTTTGGAAGAGCTAATCTGTGCAGACTG 897  
 Db 253 GLYHisAspAlaProProlalleleuThrPheGluGlnAlaAsnLeuYsgIntIleu 272  
 QY 898 AATACACATTCCTTAAGCTGGTTACTAAGCTTACTCTGTGCAAAATACACTATT 957  
 Db 273 AasnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIle 292  
 QY 958 CCTATC 963  
 Db 293 ProIle 294  
 RESULT 3  
 DDX4\_RAT  
 ID DDX4\_RAT STANDARD; PRT; 713 AA.  
 AC Q64060;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DEAD-box protein 4 (VASA homolog) (rVLSG).  
 GN DDX4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Testis;  
 RX MEDLINE=95160706; PubMed=7857296;  
 RA Komiya T., Tanigawa Y.;  
 RT "Cloning of a gene of the DEAD box protein family which is  
 RT specifically expressed in germ cells in rats.";   
 BL Biochem. Biophys. Res. Commun. 207:405-410(1995).  
 CC -1- FUNCTION: May play a role in germ cell development.  
 CC -1- TISSUE SPECIFICITY: Testis.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
 CC DDX4/VASA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; S75275; AAB33364.1; -;  
 DR HSSP; Q58083; IHV8.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC\_C; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 KM Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.  
 FT NP\_BIND 317 324  
 FT SITE 431 434  
 FT DOMAIN 58 219  
 SQ SEQUENCE 713 AA; 77955 MW; A31A4542EF7237F6 CRC64;  
 FT  
 Alignment Scores:  
 Pred. No.: 4.14e-74 Length: 713  
 Score: 82.00 Matches: 82



Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.33% Indels: 0  
 DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x DDX4\_RAT (1-713)

QY 718 CAGAGCAAAAGTGAAGCTTACATCCCTCTCCATCGAGGATGAGAGATCTCATCTTT 777  
 |||||  
 DB 225 GInGlyProlYvalThrIlyrIleProProProProGluuSpIunSpSerIlePhe 244  
 |||||

QY 778 GCACTTATCAGACAGGCACTTAACCTTCGACAAATAGACACTTCTTGTGGAGAGTCT 837  
 |||||  
 DB 245 AAlaHsTyGlnThrIlyIleAsnPhenAspIyStyRAspThrIleuValGluValSer 264  
 |||||

QY 838 GGACATGAGCACCACGACATTCGACTTGTGAGAGCAATATCTGTGACACAG 897  
 |||||  
 DB 265 GInHsAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuGlyGlnThrLeu 284  
 |||||

QY 898 AATAACACATTCCTTAAGCTGGTTACTTAAGCTTACTCTGTGCAAAATACAGATT 957  
 |||||  
 DB 285 AAsnAsnIleAlaIlyAlaIlyAlaGlyIyTrIlySleuThrProValGlnIlyStySerIle 304  
 |||||

QY 958 CCTATC 963  
 |||||

DB 305 Prolle 306

RESULT 4  
 VASA\_DROME STANDARD; PRT; 661 AA.  
 ID VASA\_DROME STANDARD; PRT; 661 AA.  
 AC P09052; Q24582; Q9V3Q8;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Vasa protein (Antigen Mab46f11).  
 GN VAS OR BG:DS00929.14 OR CG3506.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 NC NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89014721; PubMed=3140040;  
 RA Lasko P.F., Ashburner M.;  
 RT "The product of the Drosophila gene vasa is very similar to  
 RL eukaryotic initiation factor-4A.";  
 RL Nature 335:611-617(1988).  
 RN (2)  
 RP REVISIONS.  
 RA Lasko P.F.;  
 RN Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89028669; PubMed=3052853;  
 RA Hay B., Jan L.Y., Jan Y.N.;  
 RT "A protein component of Drosophila polar granules is encoded by vasa  
 RL and has extensive sequence similarity to ATP-dependent helicases.";  
 RL Cell 55:577-587(1988).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=99403001; PubMed=10471707;  
 RA Ashburner M., Mistra S., Roote J., Lewis S.E., Blazej R., Davis T.,  
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,  
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,  
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,  
 RA Celisner S., Rubin G.M.;  
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of  
 RL Drosophila melanogaster: the Adh region.";  
 RL Genetics 153:179-219(1999).  
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R., Galle R.F.,  
 RA Suton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baker E.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Bouchan M.R., Bouck J., Brokstein P., Brotier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacible J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirek Z., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).

CC - FUNCTION: THE VASA PROTEIN IS REQUIRED ONLY IN THE FEMALE GERM  
 CC LINE. IT IS IMPORTANT FOR OOCYTE FORMATION AND IN THE  
 CC SPECIFICATION OF THE POSTERIOR STRUCTURES OF THE EMBryo.  
 CC - DEVELOPMENTAL STAGE: MATERNALLY EXPRESSED (DURING Oogenesis).  
 CC FUNCTION DURING EARLY EMBRYOGENESIS.  
 CC - SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
 CC DDX4/VASA SUBFAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X12945; CAA31405.1;  
 DR EMBL: X12946; CAA31405.1; JOINED.  
 DR EMBL: W23560; AAA29013.1;  
 DR EMBL: AE003412; AAF44917.1;  
 DR EMBL: AE003646; AAF53438.1;  
 DR PIR: A31922; A31922.  
 DR PIR: S01676; S01676.  
 DR HSSP: Q58083; 1HV8.  
 DR FlyBase: FBgn0003970; vas.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR000629; DEAD\_box.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR Pfam: PF00270; DEAD; 1.  
 DR Pfam: PF00271; Helicase\_C; 1.  
 DR SMART: SM00487; DEXDC1.  
 DR SMART: SM00490; HELIC\_C; 1.  
 DR PROSITE: PS00039; DEAD\_ATP\_HELICASE; 1.

KM Developmental protein; Hydrolase; ATP-binding; Repeat; Helicase.  
FT NP\_BIND 289 ATP (BY SIMILARITY).  
FT SITE 399 DEAD BOX.  
FT DOMAIN 93 127 5 X 7 AA TANDEM REPEATS OF [FS]-R-G-G-[EO]-G-G.  
FT REPEAT 93 99 1.  
FT REPEAT 100 106 2.  
FT REPEAT 107 113 3.  
FT REPEAT 114 120 4.  
FT REPEAT 121 127 5.  
FT CONFLICT 121 127 A -> R (IN REF. 3).  
FT CONFLICT 153 165 MISSING (IN REF. 3).  
FT CONFLICT 192 192 V -> A (IN REF. 1 AND 3).  
FT CONFLICT 265 265 Y -> F (IN REF. 1).  
FT CONFLICT 322 322 V -> C (IN REF. 3).  
FT CONFLICT 452 452 F -> S (IN REF. 1).  
FT CONFLICT 582 582 R -> C (IN REF. 1).  
FT CONFLICT 594 594 D -> H (IN REF. 3).  
SQ SEQUENCE 661 AA; 72331 MW; 8617C25CCB3130B9 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,82e-14 Length: 661  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.18% Indels: 0  
DB: 1 Gaps: 0  
  
US-09-714-865-15 (1-2172) x VASA\_DROME (1-661)  
QY 973 GAGAGATTGATGCTTGCTCAACAGAGGCTGGGAGACTGGCGCTTTCTCTTA 1032  
Db 282 G|A|g|a|s|p|e|u|m|e|c|h|a|c|y|s|a|a|g|i|n|t|h|r|g|l|y|s|e|r|g|l|y|t|h|r|a|l|a|l|a|p|h|e|u|l|e|u 301  
QY 1033 CCAATTTCG 1041  
Db 302 P|r|o|l|e|u 304  
  
RESULT 5  
PL10\_MOUSE  
ID PL10\_MOUSE STANDARD; PRT; 660 AA.  
AC P16381;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Putative ATP-dependent RNA helicase PL10.  
GN D1Pasi OR PL10.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=99249320; PubMed=2720782;  
RA Leroy P., Alzari P., Sassoon D., Wolgemuth D., Fellous M.;  
RT "The protein encoded by a murine male germ cell-specific transcript  
is a putative ATP-dependent RNA helicase."  
RL Cell 57:549-559(1989).  
CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. POSSIBLE ROLE  
IN A KEY STEP OF THE SPERMATOGENIC PROCESS.  
CC -1- TISSUE SPECIFICITY: TESTIS.  
CC -1- DEVELOPMENTAL STAGE: HIGH LEVELS OF PL10 DURING THE MEIOTIC AND  
HAPLOID STAGES OF SPERMATOGENESIS.  
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC EMBL: J04847; AAA39942.1; -  
DR PIR; A32378; A32378.  
DR HSSP; Q58083; IHV8.  
DR MGD; MG1:91842; D1Pasi.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR000629; DEAD box.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00270; DEAD; 1. C1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR PROSITE; PS00039; DEAD ATP HELICASE; 1.  
KM Helicase; ATP-binding; RNA-binding; DNA-binding; Spermatogenesis.  
FT NP\_BIND 223 220 ATP (POTENTIAL).  
FT SITE 346 349 DEAD BOX.  
FT DNA\_BIND 495 514 POTENTIAL.  
FT DOMAIN 581 660 GUY/SER-RICH.  
SQ SEQUENCE 660 AA; 73140 MW; 50AD468A131AE603 CRC64;  
  
Alignment Scores:  
Pred. No.: 1,87e-13 Length: 660  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.04% Indels: 0  
DB: 1 Gaps: 0  
  
US-09-714-865-15 (1-2172) x PL10\_MOUSE (1-660)  
QY 976 CGAGATTGATGCTTGCTCAACAGAGGCTGGGAGACTGGCGCTTTCTCTTACCA 1035  
Db 217 A|g|a|s|p|e|u|m|e|c|h|a|c|y|s|a|a|g|i|n|t|h|r|g|l|y|s|e|r|g|l|y|t|h|r|a|l|a|l|a|p|h|e|u|l|e|u 236  
QY 1036 A|T|T|T|G 1041  
Db 237 I|l|e|u 238  
  
RESULT 6  
DDX3\_HUMAN  
ID DDX3\_HUMAN STANDARD; PRT; 661 AA.  
AC O00571; O15536;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE DEAD-box protein 3 (Helicase-like protein 2) (HLP2) (DEAD-box, X  
isoform).  
GN DDX3 OR DBX.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver, and Hippocampus;  
RA Chung J., Lee S.-G., Song K.;  
RT "Identification of a human homolog of a putative RNA helicase gene  
(mDEAD3) expressed in mouse erythroid cells."  
RL Korean J. Biochem. 27:193-197(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Owsianka A.M., Patel A.H.;  
RT "DEAD box putative RNA helicase."  
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98022381; PubMed=9381176;  
RA Iahn B.T., Page D.C.;  
RT "Functional coherence of the human Y chromosome."  
RL Science 278:675-680(1997).  
RN [4]  
RP SEQUENCE FROM N.A.

CC TISSUE=Skin;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RP [5]  
 RN SEQUENCE OF 1-9, AND ACETYLATION.  
 RX MEDLINE=20318637; PubMed=10859333;  
 RA Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,  
 "An N-acetylated natural ligand of human histocompatibility leukocyte  
 antigen (HLA)-B39. Classical major histocompatibility complex class I  
 proteins bind peptides with a blocked NH(2) terminus in vivo."  
 RL J. Exp. Med. 191:2083-2092(2000).  
 CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. INTERACTS  
 SPECIFICALLY WITH HEPATITIS C VIRUS CORE PROTEIN RESULTING A  
 CHANGE IN INTRACELLULAR LOCATION.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U50553; AAB95637.1; -;  
 DR EMBL; AF061337; AAC34298.1; -;  
 DR EMBL; AF000983; AAC31830.1; -;  
 DR EMBL; AF000982; AAC31829.1; -;  
 DR EMBL; BC011819; AAH11819.1; -;  
 DR HSSP; C58083; 1HV8.  
 DR Genew; HGNC:2745; DDX3.  
 DR MIM; 300160; -;  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00039; DEAD-ATP\_HELICASE; 1.  
 DR Helicase; ATP-binding; RNA-binding; DNA-binding; Acetylation.  
 DR INIT MET 0 0  
 FT NP BIND 223 230 ATP (POTENTIAL).  
 FT SITE 346 349 DEAD BOX.  
 FT DOMAIN 581 661 GLY/SER-RICH.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT CONFLICT 49 49 K -> R (IN REF. 3).  
 SQ SEQUENCE 661 AA; 73112 MW; F0F03DB8FBC00A65 CRC64;  
 Alignment Scores:  
 Pred. No.: 1,87e-13 Length: 661  
 Score: 22.00 Matches: 22  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.04% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-714-865-15 (1-2172) x DDX3\_HUMAN (1-661)  
 QY 976 CGAGATTGATGGCTGTGCTCAACAGGCTGGAGAGCTGGCTTTTCTCTACCA 1035  
 Db 217 ArgAspLeuMetAlaCysAlaGlnTrgIysGlyThrAlaAlaPheLeuPro 236  
 QY 1036 ATTTTG 1041  
 Db 237 IleLeu 238  
 RESULT 7  
 DDX3\_MOUSE STANDARD; PRT; 661 AA.  
 ID DDX3\_MOUSE STANDARD; PRT; 661 AA.  
 AC Q62167; O09060; O09143;

DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DEAD-box protein 3 (DEAD-box RNA helicase DEAD3) (mDEAD3) (Embryonic  
 DE RNA helicase) (DIPAS1 related sequence 2).  
 GN DDX3 OR DEAD3 OR ERH OR DIPAS1-RS2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; and DBA;  
 RX MEDLINE=97104282; PubMed=8948440;  
 RA Sowden J.C., Pitt W., Morrison K., Beddington R., Edwards Y.;  
 RT "The embryonic RNA helicase gene (ERH): a new member of the DEAD box  
 RT family of RNA helicases."  
 RL Biochem. J. 308:839-846(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Erythrocytopenia;  
 RX MEDLINE=94192995; PubMed=8144024;  
 RA Gee S.L., Conboy J.G.;  
 RT "Mouse erythroid cells express multiple putative RNA helicase genes  
 RT exhibiting high sequence conservation from yeast to mammals."  
 RL Gene 140:171-177(1994).  
 RN [3]  
 RP SEQUENCE OF 1-9, AND ACETYLATION.  
 RX MEDLINE=20318637; PubMed=10859333;  
 RA Yaguee J., Alvarez I., Rognan D., Ramos M., Vazquez J.,  
 "An N-acetylated natural ligand of human histocompatibility leukocyte  
 antigen (HLA)-B39. Classical major histocompatibility complex class I  
 proteins bind peptides with a blocked NH(2) terminus in vivo."  
 RL J. Exp. Med. 191:2083-2092(2000).  
 CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. IT MAY PLAY A ROLE  
 IN TRANSLATIONAL ACTIVATION OF MRNA IN THE OOCYTE AND EARLY  
 EMBRYO.  
 CC -1- TISSUE SPECIFICITY: DEVELOPMENTALLY REGULATED.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN OOCYTES. UBIQUITOUSLY FOUND IN 9  
 CC DAYS POST-CONCEPTION EMBRYO, AT LATER STAGES IT IS RESTRICTED TO  
 CC BRAIN AND KIDNEY.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z38117; CAA6261.1; -;  
 DR EMBL; L25126; AAB3630.1; -;  
 DR HSSP; C58083; 1HV8.  
 DR MGD; MG1:103064; Ddx3.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR PROSITE; PS00039; DEAD-ATP\_HELICASE; 1.  
 DR Helicase; ATP-binding; RNA-binding; DNA-binding; Acetylation.  
 DR INIT MET 0 0  
 FT NP BIND 223 230 ATP (POTENTIAL).  
 FT SITE 346 349 DEAD BOX.  
 FT DOMAIN 581 661 GLY/SER-RICH.  
 FT DOMAIN 608 615 POLY-SER.  
 FT DOMAIN 623 629 POLY-GLY.  
 FT DOMAIN 632 640 POLY-GLY.

```

FT MOD RES 1 1 ACETYLATION.
SQ SEQUENCE 661 AA; 72970 MW; A1E1PAB5D19F57B CRC64;

Alignment Scores:
Pred. No.: 1.87e-13 Length: 661
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.04% Indels: 0
DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x DDX3_MOUSE (1-661)
QY 976 CGAGATTGATGCTGTGCTCAACAGGCTGGAGACTGCGCTTTCTCTACCA 1035
DB 217 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyThrAlaAlaPheLeuPro 236

QY 1036 ATTTTG 1041
DB 237 I1e1eu 238

RESULT 8
AN3_XENLA STANDARD; PRT; 697 AA.
ID AN3_XENLA
AC P24346;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ATP-dependent RNA helicase An3.
GN AN3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9114586; PubMed=196140;
RA Gururajan R., Perry-O'Keefe H., Melton D.A., Weeks D.L.;
RT "The Xenopus localized messenger RNA An3 may encode an ATP-dependent
RT RNA helicase."
RL Nature 349:717-719(1991).
CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X57328; CAA40605.1; -.
DR PIR; S13654; S13654.
DR PIR; S29676; S29676.
DR HSSP; Q58083; 1HV8.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; ATP-binding; RNA-binding.
FT NP BIND 265 272 ATP (BY SIMILARITY).
FT SITE 388 391 DEAD BOX.
FT DOMAIN 623 697 GLY/SER-RICH.
SQ SEQUENCE 697 AA; 77302 MW; F3DD23EB60B2E2EF CRC64;

Alignment Scores:

```

```

Pred. No.: 1.85e-13 Length: 697
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.04% Indels: 0
DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x AN3_XENLA (1-697)
QY 976 CGAGATTGATGCTGTGCTCAACAGGCTGGAGACTGCGCTTTCTCTACCA 1035
DB 259 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuPro 278

QY 1036 ATTTTG 1041
DB 279 I1e1eu 280

RESULT 9
DDXY HUMAN
ID DDXY HUMAN STANDARD; PRT; 660 AA.
AC O15523;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DEAD-box protein 3, Y-chromosomal.
GN DBY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98022381; PubMed=9381176;
RA Lahn B.T., Page D.C.;
RT "Functional conference of the human Y chromosome."
RL Science 278:675-680(1997).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF000985; AAC51832.1; -.
DR EMBL; AF000984; AAC51831.1; -.
DR HSSP; Q58083; 1HV8.
DR GeneW; HGNC:2699; DBY.
DR MIM; 400010; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; ATP-binding; RNA-binding; DNA-binding.
FT NP BIND 222 229 ATP (POTENTIAL).
FT SITE 345 348 DEAD BOX.
SQ SEQUENCE 660 AA; 73094 MW; 5A67DC082AEB9CFF CRC64;

Alignment Scores:
Pred. No.: 2.06e-09 Length: 660
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.49% Indels: 0
DB: 1 Gaps: 0

```

US-09-714-865-15 (1-2172) x DDXY\_HUMAN (1-660)

Qy 988 GCTTGCTCAACAGAGCTGCGGCTTTCTCTACCAATTTG 1041  
 |||  
 Db 220 AlAcCysAlaGlnThrGlySerGlyLysThrAlaAlaLeuLeuProIleLeu 237

RESULT 10  
 ID GLH1\_CAEEL STANDARD; PRT; 763 AA.  
 AC P34689; O9TXH4;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ATP-dependent RNA helicase glh-1 (germline helicase-1).  
 GN GLH-1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peleodermidae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=94022363; PubMed=8415696;  
 RA Rouseell D.U., Bennett K.L.;  
 RT "glh-1, a germ-line putative RNA helicase from Caenorhabditis, has  
 four zinc fingers.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:9300-9304(1993).  
 RN [2]  
 RP REVISIONS TO 83-138; 275; 288 AND 398.  
 RA Rouseell D.U., McCrone J.S., Smith P.A., Gruidl M.E., Bennett K.L.;  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.  
 CC -1- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 4 CCHC-TYPE ZINC FINGERS.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 -----  
 CC EMBL: L19948; AAC27384.1; -;  
 DR PIR: A48686; A48686.  
 DR HSSP: Q58083; 1HW8.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR000629; DEAD\_box.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR001878; Znf.CCHC.  
 DR Pfam: PF00098; Zf-CCHC; 4.  
 DR Pfam: PF00270; DEAD; 1.  
 DR Pfam: PF00271; Helicase\_C\_1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR SMART: SM00490; HELIC\_C; 1.  
 DR SMART: SM00343; Znf.C2HC; 4.  
 DR PROSITE: PS00039; DEAD ATP HELICASE; 1.  
 DR PROSITE: PS50158; ZF\_CCHC; 4.  
 DR Helicase; ATP-binding; RNA-binding; Zinc-finger; Repeat.  
 KW DOMAIN 24 93  
 FT REPEAT 24 33 1.  
 FT REPEAT 34 43 2.  
 FT REPEAT 44 53 3.  
 FT REPEAT 54 63 4.  
 FT REPEAT 64 73 5.  
 FT REPEAT 74 83 6.  
 FT REPEAT 84 93 7.  
 FT DOMAIN 207 236 GLY-RICH.  
 FT ZN\_FING 158 175 CCHC-TYPE 1.  
 FT ZN\_FING 183 200 CCHC-TYPE 2.  
 FT ZN\_FING 242 259 CCHC-TYPE 3.  
 FT ZN\_FING 262 279 CCHC-TYPE 4.

FT NE\_BIND 385 392 ATP (BY SIMILARITY).  
 FT SITE 499 502 DEAD BOX.  
 SO SEQUENCE 763 AA; 79792 MW; ADB69DE286A028D6 CRC64;  
 Alignment Scores:  
 Pred. No.: 2.02e-09 Length: 763  
 Score: 18.00 Matches: 18  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.49% Indels: 0  
 DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x GLH1\_CAEEL (1-763)

Qy 985 ATGCTTGCTCAACAGAGCTGCGGCTTTCTCTACCAATT 1038  
 |||  
 Db 382 MetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaLeuLeuProIle 399

RESULT 11  
 ID DBP1\_YEAST STANDARD; PRT; 617 AA.  
 AC P24784; P20446;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable ATP-dependent RNA helicase DBP1 (Helicase CA1).  
 GN DBP1 OR YPL119C OR IPH8C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae.  
 OX NCBI\_TaxID=4932;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BY939;  
 RX MEDLINE=9131217; PubMed=1857205;  
 RA Jamieson D.J., Beggs J.D.;  
 RT "A suppressor of yeast spp81/deb1 mutations encodes a very similar  
 RT putative ATP-dependent RNA helicase.";  
 RL Mol. Microbiol. 5:805-812(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288c / AB972;  
 RX MEDLINE=97313271; PubMed=9169875;  
 RA Bussay H., Storms R.K., Ahmed A., Albertmann K., Allen E., Ansoorge W.,  
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
 RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,  
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,  
 RA Hunke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,  
 RA Komp C., Kurdi O., Lahekari D., Lew H., Lin A., Lin D., Louis E.J.,  
 RA Marache R., Messenguy F., Mewes H.-W., Mitsuhashi S., Moestl D.,  
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
 RA Patel F.X., Pohl T.M., Purnelle D., Schaefer M., Scharfe M.,  
 RA Scherens B., Schramm S., Schroeder M., Sidu A.M., Tettein H.,  
 RA Urrestarazu L.A., Ushinsky S., Vierdeleis F., Vissers S., Voss H.,  
 RA Walsh S.V., Wandut R., Wang Y., Wedler E., Wedler H., Wilmitt E.,  
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";  
 RL Nature 387:103-105(1997).  
 RN [3]  
 RP SEQUENCE OF 316-500 FROM N.A.  
 RX MEDLINE=90160368; PubMed=2406722;  
 RA Chang T.-H., Arenas U., Abelson J.;  
 RT "Identification of five putative yeast RNA helicase genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:1571-1575(1990).  
 CC -1- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE. SUPPRESSOR OF A  
 CC SPP81/DEB1 MUTATION.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDY3  
 CC SUBFAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X55993; CAA39465.1; -  
 DR EMBL; U43503; AAB68243.1; -  
 DR PIR; S16790; S16790.  
 DR PIR; A34848; A34848.  
 DR HSSP; Q58083; 1HV8.  
 DR SGD; S0006040; DBP1.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 KM Helicase; ATP-binding; RNA-binding  
 FT NP\_BIND 198 205 ATP (BY SIMILARITY).  
 FT SITE 318 321 DEAD BOX.  
 FT CONFLICT 43 44 ST -> RS (IN REF. 1).  
 FT CONFLICT 48 48 E -> K (IN REF. 1).  
 FT CONFLICT 88 88 G -> R (IN REF. 1).  
 FT CONFLICT 115 115 E -> QK (IN REF. 1).  
 FT CONFLICT 496 496 V -> I (IN REF. 3).  
 SQ SEQUENCE 617 AA; 67917 MW; 8152404B7628671D CRC64;

Alignment Scores:  
 Pred. No.: 2,14e-08 Length: 617  
 Score: 17.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.35% Indels: 0  
 DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x DBP1\_YEAST (1-617)

QY 1318 ATCAATAGTCTAGTTGGATGAAGCTGATGCGATGTTGGATATGCGTTT 1368  
 Db 312 Ileyetyrleuvallleuaspglualaaspargmetleuaspmetcylphe 328

RESULT 12  
 GLH3\_CABEL STANDARD; PRT; 720 AA.  
 AC 001836;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ATP-dependent RNA helicase glh-3 (Germline helicase-3).  
 GN GLH-3 OR B0414.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=20311358; PubMed=10851135;  
 RA Kuznicki K.A., Smith P.A., Leung-Chiu W.M., Estévez A.O., Scott H.C.,  
 Bennett K.L.;  
 "Combinatorial RNA interference indicates GLH-4 can compensate for  
 RT GLH-1; these two P granule components are critical for fertility in  
 RT C. elegans.";  
 RL Development 127:2907-2916(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Sammons L., Wohldmann P., Rohlfing T.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: PROBABLE ATP-BINDING RNA HELICASE.  
 CC -1- DEVELOPMENTAL STAGE: DURING GERM-LINE PROLIFERATION.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF079509; AAC28388.1; -  
 DR EMBL; AF003145; AAB57719.1; -  
 DR HSSP; Q58083; 1HV8.  
 DR WormRep; B0414.6; CE07736.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR001878; Znf CCHC.  
 DR Pfam; PF00098; zf-CCHC; 2.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR PRINTS; PR00939; C2HC2NFINGER.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR SMART; SM00343; ZNF CCHC; 2.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 DR PROSITE; PS0158; ZF\_CCHC; 2.  
 KM Helicase; ATP-binding; RNA-binding; Zinc-finger; Repeat.  
 FT ZN\_FING 202 219 CCHC-TYPE 1.  
 FT ZN\_FING 222 239 CCHC-TYPE 2.  
 FT NP\_BIND 342 349 ATP (BY SIMILARITY).  
 FT SITE 456 459 DEAD BOX.  
 SQ SEQUENCE 720 AA; 79728 MW; A91F49FE2B79DEAA CRC64;

Alignment Scores:  
 Pred. No.: 2,09e-08 Length: 720  
 Score: 17.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.35% Indels: 0  
 DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x GLH3\_CABEL (1-720)

QY 988 GCTTGCTCAACACAGGCTCTGGGAAGACTGCGGCTTTCTCTACCAATT 1038  
 Db 340 AAcyaAlaGlnThrIserGlySThrAlaAlaPheLeuLeuProIle 356

RESULT 13  
 DEAD\_MOUSE STANDARD; PRT; 245 AA.  
 AC 062095;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DEAD-box RNA helicase DEAD2 (mDEAD2) (DIPASI related sequence 1)  
 DE (fragment).  
 GN DEAD2 OR DIPASI-RS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Erythrocytoma;  
 RX MEDLINE=94192985; PubMed=8144024;  
 RA Gee S.L., Conboy J.G.;  
 "Mouse erythroid cells express multiple putative RNA helicase genes  
 RT exhibiting high sequence conservation from yeast to mammals.";  
 RL Gene 140:171-177(1994).

CC - FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE. IT MAY PLAY A ROLE  
 CC IN TRANSLATIONAL ACTIVATION OF MRNA IN THE OOCYTE AND EARLY  
 CC EMBRYO.  
 CC - TISSUE SPECIFICITY: FOUND IN HEART, BRAIN, LIVER, SKELETAL MUSCLE,  
 CC AND KIDNEY. LOW EXPRESSION WAS DETECTED IN LUNG.  
 CC - SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L25337; AAA5631.1; --  
 CC HSSP: Q58083; 1HV8.  
 CC DR MGDI: 99524; DIPa1-1s1.  
 CC DR InterPro: IPR001410; DEAD.  
 CC DR InterPro: IPR000629; DEAD\_box.  
 CC DR InterPro: IPR001650; Helicase\_C.  
 CC DR Pfam: PF00270; DEAD; 1.  
 CC DR Pfam: PF00271; Helicase\_C; 1.  
 CC DR SMART: SM00490; HELIC; 1.  
 CC DR PROSITE: PS00039; DEAD ATP HELICASE; 1.  
 CC DR Helicase; ATP-binding; RNA-binding.  
 CC FT NON\_TER 1 1  
 CC FT SITE 64 67 DEAD BOX.  
 CC FT NOM\_TER 245 245  
 CC SQ SEQUENCE 245 AA; 27978 MW; 276FC43820B3EDDA CRC64;  
 CC  
 CC Alignment Scores:  
 CC Pred. No.: 2,52e-07 Length: 245  
 CC Score: 16.00 Matches: 16  
 CC Percent Similarity: 100.00% Conservative: 0  
 CC Best Local Similarity: 100.00% Mismatches: 0  
 CC Query Match: 2.21% Indels: 0  
 CC DB: 1 Gaps: 0  
 CC  
 CC US-09-714-865-15 (1-2172) x DEAD\_MOUSE (1-245)  
 CC QY 1321 AATACTTGAATTTGGATGAAGCTGATCGCATGTTGATGCGTTT 1368  
 CC Db 59 LysTyrLeuValLeuaspGluAlaaspArgMetLeuaspMetClyPhe 74  
 CC  
 CC RESULT 14  
 CC DED1\_YEAST  
 CC ID DED1\_YEAST STANDARD; PRT; 604 AA.  
 CC AC P06634;  
 CC DT 01-JAN-1988 (Rel. 06, Created)  
 CC DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Probable ATP-dependent RNA helicase DED1.  
 CC GN DED1 OR SPB1 OR YOR204W.  
 CC OS Saccharomyces cerevisiae (Baker's Yeast).  
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 CC OX NCBI\_TaxID=4932;  
 CC RN (1)  
 CC RP SEQUENCE FROM N.A.  
 CC RX STRAIN=A364A X H79-20.3;  
 CC MEDLINE=91141585; PubMed=1996139;  
 CC RA Jamieson D.J., Rahe B., Pringle J., Beggs J.D.,  
 CC RT "A suppressor of a yeast splicing mutation (prp8-1) encodes a  
 CC RT putative ATP-dependent RNA helicase.";  
 CC RL Nature 349:715-717(1991).  
 CC [2]  
 CC RN SEQUENCE FROM N.A.  
 CC RA Hughes B., Pohl T.M.,  
 CC RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC RN [3]  
 CC RP SEQUENCE OF 1-112 FROM N.A.

RX MEDLINE=86093663; PubMed=3001645;  
 RA Struhl K.;  
 RT "Nucleotide sequence and transcriptional mapping of the yeast  
 RT prp8-hist3-ded1 gene region.";  
 RL Nucleic Acids Res. 13:887-860(1985).  
 RN (4)  
 CC CHARACTERIZATION.  
 CC RP MEDLINE=97289744; PubMed=9144215;  
 CC RA de la Cruz J., Iost I., Kressler D., Linder P.;  
 CC RT "The prp8 and ded1 proteins have antagonistic roles in eIF4E-dependent  
 CC RT translation in Saccharomyces cerevisiae.";  
 CC RL Proc. Natl. Acad. Sci. U.S.A. 94:5201-5206(1997).  
 CC - FUNCTION: POSITIVELY INVOLVED IN THE INITIATION OF PROTEIN  
 CC SYNTHESIS. PROBABLE ATP-BINDING RNA HELICASE. ESSENTIAL FOR CELL  
 CC VIABILITY. SUPPRESSOR OF A PRP8 MUTATION. MAY PLAY A ROLE IN MRNA  
 CC SPLICING.  
 CC - SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X57278; CAA40546.1; --  
 CC DR EMBL: 275110; CAA99419.1; --  
 CC DR EMBL: X03245; CAA27004.1; --  
 CC DR PIR: S13653; S13653.  
 CC DR PIR: S07683; S07683.  
 CC DR HSSP: Q58083; 1HV8.  
 CC DR SGD: S0005730; DED1.  
 CC DR InterPro: IPR001410; DEAD.  
 CC DR InterPro: IPR000629; DEAD\_box.  
 CC DR InterPro: IPR001650; Helicase\_C.  
 CC DR Pfam: PF00270; DEAD; 1.  
 CC DR Pfam: PF00271; Helicase\_C; 1.  
 CC DR SMART: SM00487; DEXDC; 1.  
 CC DR SMART: SM00490; HELIC; 1.  
 CC DR PROSITE: PS00039; DEAD ATP HELICASE; 1.  
 CC DR Helicase; ATP-binding; RNA-binding.  
 CC FT NP BIND 186 193 ATP (BY SIMILARITY).  
 CC FT SITE 306 309 DEAD BOX.  
 CC FT CONFLICT 37 37 S->M (IN REF. 2).  
 CC SQ SEQUENCE 604 AA; 65553 MW; B6722D94C03BFA4B CRC64;  
 CC  
 CC Alignment Scores:  
 CC Pred. No.: 2.2e-07 Length: 604  
 CC Score: 16.00 Matches: 16  
 CC Percent Similarity: 100.00% Conservative: 0  
 CC Best Local Similarity: 100.00% Mismatches: 0  
 CC Query Match: 2.21% Indels: 0  
 CC DB: 1 Gaps: 0  
 CC  
 CC US-09-714-865-15 (1-2172) x DED1\_YEAST (1-604)  
 CC QY 1321 AATACTTGAATTTGGATGAAGCTGATCGCATGTTGATGCGTTT 1368  
 CC Db 301 LysTyrLeuValLeuaspGluAlaaspArgMetLeuaspMetClyPhe 316  
 CC  
 CC RESULT 15  
 CC DED1\_SCHPO  
 CC ID DED1\_SCHPO STANDARD; PRT; 636 AA.  
 CC AC O13370; O59857;  
 CC DT 15-JUN-2002 (Rel. 41, Created)  
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE ATP-dependent RNA helicase ded1.  
 CC GN DED1 OR SUM3 OR DEPI OR MOC2 OR SPEC1795.11.  
 CC OS Schizosaccharomyces pombe (fission yeast).  
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

CC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 CC Schizosaccharomyces.  
 CX NCBI\_TaxID=4896;  
 RN  
 RP SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE FORMS.  
 RX MEDLINE=21659741; PubMed=1711540;  
 RA Liu H.-Y., Nefsky B.S., Maitworth N.C.;  
 RT "The ded1 DEAD box helicase interacts with Chk1 and Cdc2.";  
 RL J. Biol. Chem. 277:2637-2643(2002).  
 RN  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RX MEDLINE=20191909; PubMed=10725227;  
 RA Graillet B., Kearsey S.E., Lenhard M., Carlson C.R., Nurse P.,  
 RA Boye E., Labib K.;  
 RL "A fission yeast general translation factor reveals links between  
 RT protein synthesis and cell cycle controls.";  
 RL J. Cell Sci. 113:1447-1458(2000).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99326140; PubMed=10395922;  
 RA Kawamukai M.;  
 RL "Isolation of a novel gene, mcc2, encoding a putative RNA helicase as  
 RT a suppressor of sterile strains in Schizosaccharomyces pombe.";  
 RL Biochim. Biophys. Acta 1446:93-101(1999).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99054689; PubMed=9832516;  
 RA Forbes K.C., Humphrey T., Enoch T.;  
 RT "Suppressors of cdc25p overexpression identify two pathways that  
 RL influence the G2/M checkpoint in fission yeast.";  
 RL Genetics 150:1361-1375(1998).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21848401; PubMed=11895360;  
 RL STRAIN=972;  
 RA Wood V., Gwilliam R., Rajandream A.A., Lyne M., Lyne R., Stewart A.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds R., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymourez B.,  
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goiteanu A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motlier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,  
 RA Ducas R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Renuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,  
 RA Cerruti L., Lowe T., McCombie M.R., Paulsen I., Fothergill J.,  
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC  
 CC -1- FUNCTION: Positively involved in the initiation of protein  
 CC synthesis. Probable ATP-binding RNA helicase. Essential for cell  
 CC viability. May play a role in mRNA splicing. Inactivation of ded1  
 CC blocks mitotic cell cycle progression at G1 and G2/M.  
 CC -1- SUBUNIT: Interacts with chk1, which is required for cell cycle  
 CC arrest following DNA damage.  
 CC -1- MISCELLANEOUS: A different form of ded1 has been identified via  
 CC SOS-PAGE studies. It is uncertain how this form arises in vivo.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DDX3  
 CC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF025536; AAC04893.1; -
DR EMBL; AF084222; AAC34121.1; -
DR EMBL; AB012389; BAA25324.1; -
DR EMBL; AJ237697; CAB40192.1; -
DR EMBL; AL022598; CAA18646.1; -
DR HSSP; Q58083; IHV8.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_ATP_helicase.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; ATP-binding; RNA-binding; Cell cycle; Mitosis.
FT NP_BIND 213 220 ATP (POTENTIAL).
FT SITE 336 339 DEAD BOX.
FT CONFLICT 41 41 A -> R (IN REF. 3).
SQ SEQUENCE 636 AA; 69758 MW; 094630A41A3C26F1 CRC64;

Alignment Scores:
Pred. No.: 2,18e-07 Length: 636
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Percent Local Similarity: 100.00% Mismatches: 0
Query Match: 2.21% Indels: 0
DB: 1 Gaps: 0

US-09-714-865-15 (1-2172) x DED1_SCHPO (1-636)
QY 973 GGACGAGATTGATGCTTGCTCAAAACAGGCTTGGAGACTGCG 1020
Db 206 G1YArgspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 221

Search completed: June 10, 2003, 17:05:53
Job time : 53.5 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:57:41 ; Search time 38 Seconds  
(without alignments)  
10989.682 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 724  
Sequence: 1 atcg99gatgaagatcg99ga.....tagatgatgagcatcg9gat 2172

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 1000 summaries

Command line parameters:

-MODE=frame+ n2p model -DEV=xlp  
-Q/cgr2.1/USFTO\_spool/US09714865/runat\_05062003\_111833\_26595/app\_query.fasta.1.2311  
-DB=PIR\_73 -GPM=Iaetan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=1000  
-DOCAIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORMext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09714865 @CGN 1.1 55 @runat\_05062003\_111833\_26595 -NCPU=6 -ICPU=3  
-NO\_MMAP -LANG=OTHER -NEG\_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: PIR\_73:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	522	72.1	635	2	T46407
2	82	11.3	713	2	JC2534
3	64	8.8	637	2	I49638
4	24	3.3	700	2	I51235
5	23	3.2	661	2	A58768
6	22	3.0	660	2	A32378
7	22	3.0	662	1	I84741
8	19	2.2	697	1	S13654
9	19	2.6	646	2	T45677
10	18	2.5	604	2	C87818
11	18	2.5	604	2	T15132
12	18	2.5	633	2	H84854
13	17	2.3	603	2	T45671
14	17	2.3	617	2	S62003

15	17	2.3	688	2	T48796	probable ATP-depen
16	17	2.3	720	2	T15231	germline RNA helic
17	16	2.2	604	2	S13653	ATP-dependent RNA
18	16	2.2	636	2	T43543	probable ATP-depen
19	16	2.2	707	2	A46686	probable RNA helic
20	15	2.1	263	2	T51349	RNA helicase RH30
21	15	2.1	501	2	B96593	probable ethylene-
22	15	2.1	546	1	S13757	RNA helicase DBP2
23	15	2.1	550	1	S14048	RNA helicase dbp2
24	15	2.1	561	2	T22917	probable ATP-depen
25	15	2.1	575	1	S11485	RNA helicase - fru
26	15	2.1	607	1	S42639	ATP-dependent RNA
27	15	2.1	614	1	I48385	RNA helicase TN22
28	15	2.1	619	2	JC1087	RNA helicase, ATP-
29	15	2.1	614	2	T52137	ATP-dependent DEAD
30	15	2.1	650	2	S72367	ATP-dependent RNA
31	15	2.1	713	2	T48634	DRH1 DEAD box prot
32	14	1.9	422	2	H82418	ATP-dependent RNA
33	14	1.9	423	2	F82295	ATP-dependent RNA
34	14	1.9	425	2	T46969	DEAD box RNA helic
35	14	1.9	426	2	AF2395	ATP-dependent RNA
36	14	1.9	446	2	E82835	ATP-dependent RNA
37	14	1.9	446	2	A83213	probable ATP-depen
38	14	1.9	451	2	E82488	ATP-dependent RNA
39	14	1.9	454	2	E64816	probable ATP-depen
40	14	1.9	455	2	C90738	probable ATP-depen
41	14	1.9	455	2	D85588	probable ATP-depen
42	14	1.9	455	2	A10599	probable ATP-depen
43	14	1.9	458	2	B83588	RNA helicase DbpA
44	14	1.9	460	2	AF0216	ATP-dependent RNA
45	14	1.9	462	2	E81850	ATP-dependent RNA
46	14	1.9	462	2	G81850	probable ATP-depen
47	14	1.9	535	2	AH3368	ATP-dependent RNA
48	14	1.9	639	2	D83591	probable ATP-depen
49	13	1.8	368	2	F69531	ATP-dependent RNA
50	13	1.8	397	2	C83163	ATP-dependent RNA
51	13	1.8	397	2	D82203	ATP-dependent RNA
52	13	1.8	411	2	T51739	RNA helicase RH5 (
53	13	1.8	457	2	F81857	probable ATP-depen
54	13	1.8	457	2	H81085	ATP-dependent RNA
55	13	1.8	504	2	T27176	probable ATP-depen
56	13	1.8	537	2	A86444	probable RNA helic
57	12	1.7	441	2	AD0330	ATP-dependent RNA
58	12	1.7	444	2	AD0830	ATP-dependent RNA
59	12	1.7	444	2	B91059	ATP-dependent RNA
60	12	1.7	444	2	G65035	ATP-dependent RNA
61	12	1.7	542	2	T05988	hypothetical prote
62	12	1.7	578	2	S67386	probable ATP-depen
63	12	1.7	701	2	T15942	hypothetical prote
64	12	1.7	1032	2	A57514	RNA helicase HEL11
65	11	1.5	343	2	A99353	ATP-dependent RNA
66	11	1.5	360	2	G90370	ATP-dependent RNA
67	11	1.5	444	2	H85903	ATP-dependent RNA
68	11	1.5	449	2	F83151	ATP-dependent RNA
69	11	1.5	452	2	C82506	probable ATP-depen
70	11	1.5	457	2	G90869	ATP-dependent RNA
71	11	1.5	457	2	B85749	ATP-dependent RNA
72	11	1.5	457	2	B64684	ATP-dependent RNA
73	11	1.5	482	2	AE3381	ATP-dependent RNA
74	11	1.5	493	2	G86297	F309.8 protein - A
75	11	1.5	503	2	A97581	ATP-dependent RNA
76	11	1.5	503	2	AG2801	ATP-dependent RNA
77	11	1.5	517	2	B87478	ATP-dependent RNA
78	11	1.5	523	1	S30805	hypothetical prote
79	11	1.5	543	2	S46713	probable RNA helic
80	11	1.5	601	2	D84973	ATP-dependent RNA
81	11	1.5	613	2	F64056	ATP-dependent RNA
82	11	1.5	646	2	C91134	probable ATP-depen
83	11	1.5	646	2	F85979	inducible ATP-inde
84	11	1.5	646	2	AB0901	ATP-dependent RNA
85	11	1.5	646	2	F65106	probable ATP-depen
86	11	1.5	659	2	JX0314	DEAD box protein -
87	11	1.5	662	2	T41512	us snrnp-like RNA

88	11	1.5	664	2	A10423	cold-shock dead-bo	161	8	1.1	71	2	T28271	ORF MSV109 hypothe
89	11	1.5	1032	2	G89427	protein T08D2.3 (1	c 162	8	1.1	83	2	S24211	Ig kappa chain V r
90	11	1.5	1156	2	T43336	germline RNA helic	c 163	8	1.1	91	2	S24216	Ig kappa chain V r
91	11	1.5	1172	2	T32759	hypothetical prote	c 164	8	1.1	92	2	A31409	actin - California
92	10	1.4	167	2	T51345	RNA helicase RH20	c 165	8	1.1	100	2	P90387	hypothetical prote
93	10	1.4	224	2	T51742	RNA helicase RH11	c 166	8	1.1	101	2	A33730	Ig kappa chain V r
94	10	1.4	240	2	T46269	hypothetical prote	c 167	8	1.1	112	2	A36259	Ig kappa chain V r
95	10	1.4	376	2	T51342	RNA helicase RH10	c 168	8	1.1	113	2	F30560	Ig kappa chain V r
96	10	1.4	418	2	E64100	probable ATP-depen	c 169	8	1.1	115	1	K2HUCM	Ig kappa chain V-I
97	10	1.4	421	2	A91218	probable ATP-depen	c 170	8	1.1	121	2	S24205	Ig kappa chain V r
98	10	1.4	421	2	B86064	probable ATP-depen	c 171	8	1.1	122	2	AD2072	hypothetical prote
99	10	1.4	421	2	AG0922	probable ATP-depen	c 172	8	1.1	129	2	G84768	hypothetical prote
100	10	1.4	421	2	G65181	th1b protein - Esc	c 173	8	1.1	130	2	S40321	Ig kappa chain - h
101	10	1.4	425	1	H69124	translation initia	c 174	8	1.1	132	2	C32513	Ig kappa chain pre
102	10	1.4	428	1	AE0471	probable DEAD-box	c 175	8	1.1	143	2	E64602	ribonuclease H - H
103	10	1.4	438	2	C82340	ATP-dependent RNA	c 176	8	1.1	159	2	T31598	hypothetical prote
104	10	1.4	460	2	B82060	ATP-dependent RNA	c 177	8	1.1	174	2	T29783	hypothetical prote
105	10	1.4	470	2	AE1584	ATP-dependent RNA	c 178	8	1.1	188	2	T02247	hypothetical prote
106	10	1.4	470	2	AF1230	ATP-dependent RNA	c 179	8	1.1	259	2	T51745	Ig kappa chain - h
107	10	1.4	502	2	C97640	probable ATP-depen	c 180	8	1.1	265	2	S46534	Ig kappa chain pre
108	10	1.4	502	2	AD2863	dead-box ATP-depen	c 181	8	1.1	267	1	CTPGP	ribonuclease H - H
109	10	1.4	503	2	F95988	probable ATP-depen	c 182	8	1.1	267	2	AD1835	hypothetical prote
110	10	1.4	511	2	D69772	ATP-dependent RNA	c 183	8	1.1	268	2	T51678	hypothetical prote
111	10	1.4	513	2	AD1959	ATP-dependent RNA	c 184	8	1.1	271	2	G01790	myb-related transcr
112	10	1.4	516	2	AC1540	ATP-dependent RNA	c 185	8	1.1	274	2	B70695	probable thio1-spe
113	10	1.4	520	2	AB1183	ATP-dependent RNA	c 186	8	1.1	290	2	T05009	hypothetical prote
114	10	1.4	524	2	G95184	ATP-dependent RNA	c 187	8	1.1	296	2	A47318	RNA-binding protei
115	10	1.4	524	2	G98051	conserved hypochet	c 188	8	1.1	296	2	T53142	gene Merc protein
116	10	1.4	527	2	D70595	probable ATP-depen	c 189	8	1.1	297	2	C48210	ezyl-1 protein (C10
117	10	1.4	528	2	D97270	ATP-dependent RNA	c 190	8	1.1	304	2	D84652	probable casein Ki
118	10	1.4	539	2	H83947	ATP-dependent RNA	c 191	8	1.1	320	2	C85440	myb-related protei
119	10	1.4	544	2	E87010	probable ATP-depen	c 192	8	1.1	320	2	S09208	chorion protein s3
120	10	1.4	547	2	F86668	ATP-dependent RNA	c 193	8	1.1	323	2	C83901	riboflavin kinase
121	10	1.4	563	2	E70752	probable dead prot	c 194	8	1.1	323	2	AH2139	hypothetical prote
122	10	1.4	564	2	S59649	ATP-dependent RNA	c 195	8	1.1	328	2	B45543	malosaccharide ut
123	10	1.4	567	2	A83292	probable ATP-depen	c 196	8	1.1	328	2	B95247	malose operon tra
124	10	1.4	610	2	G86407	hypothetical prote	c 197	8	1.1	328	2	G98111	malose operon tra
125	10	1.4	614	2	G82830	ATP-dependent RNA	c 198	8	1.1	341	2	B75374	conserved hypochet
126	10	1.4	646	2	H88637	protein F53H1.1 (1	c 199	8	1.1	345	2	C95873	probable malate de
127	10	1.4	748	2	H84913	probable ATP-depen	c 200	8	1.1	345	2	AH3131	malate dehydrogena
128	10	1.4	845	2	D96799	hypothetical prote	c 201	8	1.1	345	2	AD3635	malate dehydrogena
129	10	1.4	1166	2	H86341	hypothetical prote	c 202	8	1.1	346	2	C96156	probable L-malate
130	9	1.3	162	2	C85356	glycine-rich prote	c 203	8	1.1	351	2	C82755	conserved hypochet
131	9	1.3	168	2	A03864	hypothetical 17.7K	c 204	8	1.1	360	2	T51343	RNA helicase RH18
132	9	1.3	215	2	T49743	probable rrm-type	c 205	8	1.1	367	1	E64383	translation initia
133	9	1.3	228	2	T01350	hypothetical prote	c 206	8	1.1	374	2	E96996	ATP-dependent RNA
134	9	1.3	323	2	JC4774	fiber protein - hu	c 207	8	1.1	379	2	E88109	protein T24E12.8 (
135	9	1.2	376	2	C69813	RNA helicase homol	c 208	8	1.1	381	2	H90136	RNA helicase (limp)
136	9	1.2	393	2	T49578	hypothetical prote	c 209	8	1.1	382	2	S09139	coil intron protei
137	9	1.2	431	2	S48908	helicase homolog -	c 210	8	1.1	399	2	T46259	hypothetical prote
138	9	1.2	439	2	H64066	ATP-dependent RNA	c 211	8	1.1	407	2	A71673	probable ATP-depen
139	9	1.2	439	2	JC2118	DEAD-box RNA helic	c 212	8	1.1	414	2	D97827	ATP-dependent RNA
140	9	1.2	441	2	AB1280	ATP-dependent RNA	c 213	8	1.1	415	2	D96759	probable serine ca
141	9	1.2	442	2	AH1661	ATP-dependent RNA	c 214	8	1.1	423	2	G96554	hypothetical prote
142	9	1.3	453	2	B48210	ezyl-1 protein prec	c 215	8	1.1	428	2	H82390	ATP-dependent RNA
143	9	1.2	481	2	S62423	ATP-dependent RNA	c 216	8	1.1	438	2	B96759	protein serine car
144	9	1.2	481	2	C97238	ATP-dependent RNA	c 217	8	1.1	441	2	C96759	protein serine car
145	9	1.3	508	2	A32225	nerve growth facto	c 218	8	1.1	451	2	T51744	RNA helicase RH13
146	9	1.3	543	2	A41211	early growth respo	c 219	8	1.1	453	2	S74698	UDP-glucose dehydr
147	9	1.2	582	2	S53814	DEAD box protein -	c 220	8	1.1	462	4	S33798	FUS/CHOP mutant fu
148	9	1.2	606	2	S14942	RNA helicase SPB4	c 221	8	1.1	464	2	A56600	intermediate filam
149	9	1.2	733	2	H84748	hypothetical prote	c 222	8	1.1	465	2	E96737	probable DEAD/DEAH
150	9	1.2	746	2	T29584	hypothetical prote	c 223	8	1.1	469	1	I46076	DNA-directed DNA p
151	9	1.2	752	2	S64750	probable ATP-depen	c 224	8	1.1	474	2	E85175	ATP-dependent RNA
152	9	1.2	754	2	S62561	ATP dependent RNA	c 225	8	1.1	476	2	H87352	hypothetical prote
153	9	1.3	788	2	US0747	regulatory protein	c 226	8	1.1	477	2	T01202	probable RNA helic
154	9	1.2	789	2	T51310	RNA helicase RH28	c 227	8	1.1	478	2	I49731	ATP-dependent RNA
155	9	1.3	809	2	S55344	outer envelope mem	c 228	8	1.1	479	2	E69613	atp-dependent rna
156	9	1.2	811	2	T19974	hypothetical prote	c 229	8	1.1	480	2	E90520	atp-dependent rna
157	9	1.2	819	2	T08745	probable RNA helic	c 230	8	1.1	482	2	S66920	probable RNA helic
158	9	1.2	848	2	S59645	probable atp-depen	c 231	8	1.1	489	2	S40731	ATP-dependent RNA
159	9	1.2	848	2	S67564	hypothetical prote	c 232	8	1.1	491	2	B71957	ATP-dependent RNA
160	9	1.2	1212	2	C82834	DNA polymerase III	c 233	8	1.1	492	2	G64550	ATP-dependent RNA

234	8	1.1	492	2	S76612	C 307	7	1.0	63	2	T12121	NADH dehydrogenase
C 235	8	1.1	496	2	F75257	C 308	7	1.0	64	2	A64247	ribosomal protein
236	8	1.1	496	2	T04020	C 309	7	1.0	64	2	E87083	50S ribosomal
237	8	1.1	498	2	T51743	C 310	7	1.0	65	2	T25646	hypothetical prote
238	8	1.1	498	2	T47916	C 311	7	1.0	65	2	B89948	hypothetical prote
239	8	1.1	499	2	T01230	C 312	7	1.0	68	2	T29417	hypothetical prote
240	8	1.1	503	2	T39375	C 313	7	1.0	72	2	I53411	acetylcholinestera
241	8	1.1	505	2	T51741	C 314	7	1.0	73	2	S78280	conserved hypotet
242	8	1.1	505	2	S47451	C 315	7	1.0	73	2	AH2788	hypothetical prote
243	8	1.1	506	2	G90000	C 316	7	1.0	83	2	A97568	hypothetical prote
C 244	8	1.1	526	1	S33799	C 317	7	1.0	89	2	B25155	Ig kappa chain V r
245	8	1.1	527	2	S38329	C 318	7	1.0	90	2	G71368	conserved hypotet
246	8	1.1	528	2	T02466	C 319	7	1.0	91	1	WSWL13	B5 protein - human
247	8	1.1	544	2	D82526	C 320	7	1.0	91	2	F81996	hypothetical prote
248	8	1.1	546	2	T14604	C 321	7	1.0	92	2	F71868	hypothetical prote
249	8	1.1	553	2	D96684	C 322	7	1.0	95	2	F41298	kinesin-like prote
250	8	1.1	566	2	S53813	C 323	7	1.0	96	2	A57483	3-mercaptopyruvate
C 251	8	1.1	569	1	KRMSE1	C 324	7	1.0	98	2	S53815	DEAD box protein -
C 252	8	1.1	570	2	S07330	C 325	7	1.0	99	2	AB2242	hypothetical prote
253	8	1.1	571	2	T33113	C 326	7	1.0	104	1	G69482	transcription regu
254	8	1.1	572	2	G02845	C 327	7	1.0	104	2	E69767	50S ribosomal prot
255	8	1.1	578	2	S62574	C 328	7	1.0	105	2	B90021	hypothetical prote
256	8	1.1	588	2	A39624	C 329	7	1.0	109	2	G87661	hypothetical prote
257	8	1.1	595	2	T41007	C 330	7	1.0	110	2	F84670	hypothetical prote
258	8	1.1	604	2	T41249	C 331	7	1.0	112	2	D90193	conserved hypotet
259	8	1.1	606	2	T39966	C 332	7	1.0	112	4	S59333	hypothetical prote
260	8	1.1	610	2	S71758	C 333	7	1.0	116	2	S57106	probable membrane
261	8	1.1	613	2	T24662	C 334	7	1.0	116	2	A72685	hypothetical prote
262	8	1.1	615	2	F98217	C 335	7	1.0	117	2	T11933	NADH2 dehydrogenas
263	8	1.1	615	2	AE3069	C 336	7	1.0	118	2	T16510	hypothetical prote
264	8	1.1	626	2	T10237	C 337	7	1.0	120	2	T19392	hypothetical prote
265	8	1.1	626	2	AB3480	C 338	7	1.0	121	2	H82792	conserved hypotet
266	8	1.1	633	2	S69734	C 339	7	1.0	124	2	T04505	hypothetical prote
267	8	1.1	638	2	T24661	C 340	7	1.0	125	2	E90762	probable regulator
C 268	8	1.1	643	1	KRMU2	C 341	7	1.0	126	2	S40341	Ig kappa chain -h
269	8	1.1	648	2	T40504	C 342	7	1.0	126	2	T00610	hypothetical prote
C 270	8	1.1	659	2	T02838	C 343	7	1.0	127	2	F72462	hypothetical prote
271	8	1.1	663	2	C82415	C 344	7	1.0	130	2	T34729	probable gas vesic
272	8	1.1	683	2	D71433	C 345	7	1.0	132	2	A86937	probable dihydroene
273	8	1.1	685	2	T09159	C 346	7	1.0	132	2	F85625	hypothetical prote
274	8	1.1	709	2	T39930	C 347	7	1.0	133	2	F75552	ribosomal protein
275	8	1.1	742	2	S38093	C 348	7	1.0	133	2	S45752	hypothetical prote
276	8	1.1	748	2	T51738	C 349	7	1.0	136	2	E84405	hypothetical prote
277	8	1.1	752	2	D95246	C 350	7	1.0	139	2	F97438	hypothetical prote
278	8	1.1	752	2	A98111	C 351	7	1.0	140	2	D90039	hypothetical prote
279	8	1.1	770	2	S56805	C 352	7	1.0	143	2	T34647	probable integral
280	8	1.1	773	2	S46011	C 353	7	1.0	143	2	T15722	hypothetical prote
C 281	8	1.1	787	2	C84898	C 354	7	1.0	144	2	B82623	TolR protein Xf189
C 282	8	1.1	801	2	A37353	C 355	7	1.0	144	2	F87452	transcription regu
C 283	8	1.1	801	2	PC6010	C 356	7	1.0	146	2	D95244	transcription regu
C 284	8	1.1	805	2	A56199	C 357	7	1.0	147	2	A53180	ribonuclease PL3 (
C 285	8	1.1	810	2	T48835	C 358	7	1.0	147	2	T52489	ribonuclease 4 (EC
C 286	8	1.1	823	2	A36378	C 359	7	1.0	149	2	S20316	synaptophysin homo
C 287	8	1.1	853	2	A95269	C 360	7	1.0	150	2	F84976	3-dehydroquininate d
C 288	8	1.1	854	1	MMBP12	C 361	7	1.0	150	2	D48835	kinesin-like prote
C 289	8	1.1	858	2	A46613	C 362	7	1.0	150	2	C86224	hypothetical prote
C 290	8	1.1	863	2	B69527	C 363	7	1.0	153	2	AH2203	bacterioferritin c
C 291	8	1.1	873	2	H53225	C 364	7	1.0	153	2	B72721	hypothetical prote
292	8	1.1	920	2	T41050	C 365	7	1.0	155	2	B70370	hypothetical prote
293	8	1.1	971	2	S50912	C 366	7	1.0	155	2	B97207	probable acetyltra
294	8	1.1	1022	2	T24663	C 367	7	1.0	156	2	D70827	ribosomal protein
C 295	8	1.1	1295	2	A32901	C 368	7	1.0	156	2	S31149	ribosomal protein
C 296	8	1.1	1585	2	T31611	C 369	7	1.0	156	2	A87144	30S ribosomal prot
C 297	8	1.1	1660	2	A70869	C 370	7	1.0	157	1	R3CMS7	ribosomal protein
C 298	8	1.1	2083	2	T42721	C 371	7	1.0	157	2	B26956	ribosomal protein
C 299	8	1.1	2142	2	B35098	C 372	7	1.0	157	2	H72065	ribosomal protein
C 300	8	1.1	2175	1	S03170	C 373	7	1.0	157	2	H81672	homeotic protein H
C 301	7	1.0	28	2	S58389	C 374	7	1.0	157	2	S25479	hypothetical prote
C 302	7	1.0	42	2	D85604	C 375	7	1.0	157	2	B84698	transcription fac
C 303	7	1.0	45	2	H70249	C 376	7	1.0	159	2	T05279	hypothetical prote
C 304	7	1.0	51	2	S64676	C 377	7	1.0	159	2	S37766	hypothetical prote
C 305	7	1.0	63	2	G85624	C 378	7	1.0	159	2	C75446	MutT/nudix family
306	7	1.0	63	2	B90761	C 379	7	1.0	159	2	S61040	probable membrane

C 380	7	1.0	161	2	AC0439	conserved hypotet	453	7	1.0	228	2	AE0325	urease accessory p
C 381	7	1.0	161	2	AG1332	hypothetical prote	454	7	1.0	229	1	C45345	vif protein - capr
C 382	7	1.0	162	2	A45362	corticosterin pre	455	7	1.0	229	1	J01932	vif protein - ovin
C 383	7	1.0	162	2	T33127	hypothetical prote	456	7	1.0	230	1	Q0LJVS	vif protein - Maed
C 384	7	1.0	165	1	RMEC18	primosomal operon	457	7	1.0	231	2	U01163	vif protein - Maed
C 385	7	1.0	165	2	S01510	NADH dehydrogenas	C 458	7	1.0	232	2	H87113	conserved hypotet
C 386	7	1.0	165	2	H91293	probable glycoprot	C 459	7	1.0	232	2	C85585	unknown protein en
C 387	7	1.0	165	2	C86135	probable glycoprot	C 460	7	1.0	232	2	B90735	hypothetical prote
C 388	7	1.0	166	2	C65559	S7 ribosomal prote	C 461	7	1.0	232	2	J00886	hypothetical prote
C 389	7	1.0	167	2	T01155	hypothetical prote	C 462	7	1.0	233	2	B90088	similar to proteas
C 390	7	1.0	167	2	A82641	hypothetical prote	C 463	7	1.0	234	2	AB0505	fimbril chaprone
C 391	7	1.0	168	2	C75425	hypothetical prote	C 464	7	1.0	234	2	E72252	transactin initia
C 392	7	1.0	169	2	T26271	hypothetical prote	C 465	7	1.0	234	2	T49448	hypothetical prote
C 393	7	1.0	171	1	LMYCI	H+-transporting tw	C 466	7	1.0	234	2	G95989	hypothetical glyci
C 394	7	1.0	172	2	B98109	hypothetical prote	C 467	7	1.0	235	2	I40230	hypothetical prote
C 395	7	1.0	174	2	F83514	conserved hypotet	C 468	7	1.0	237	2	S55009	H+-transporting tw
C 396	7	1.0	174	2	T29067	probable ATP/GTP b	C 469	7	1.0	237	2	A98357	hypothetical prote
C 397	7	1.0	174	2	T29071	probable ATP/GTP b	C 470	7	1.0	237	2	AD2925	hypothetical prote
C 398	7	1.0	175	2	AC0853	hypothetical prote	C 471	7	1.0	238	2	T07955	H+-transporting tw
C 399	7	1.0	177	2	T00789	ubiquitin-protein	C 472	7	1.0	239	2	C82870	hypothetical prote
C 400	7	1.0	177	2	H90631	probable fibrinai	C 473	7	1.0	239	2	A97358	glucose-inhibited
C 401	7	1.0	177	2	G85482	probable type-1 fi	C 474	7	1.0	240	2	E69004	hypothetical prote
C 402	7	1.0	178	1	R3177	ribosomal protein	C 475	7	1.0	240	2	AD1552	transcription regu
C 403	7	1.0	178	2	A82606	conjugal transfer	C 476	7	1.0	240	2	AF1194	transcription regu
C 404	7	1.0	181	2	T51352	RNA helicase RH27	C 477	7	1.0	241	2	AH3543	gtp-binding protei
C 405	7	1.0	181	2	T35858	probable ATP/GTP-b	C 478	7	1.0	242	2	C82575	phage-related prot
C 406	7	1.0	182	2	A46436	myofibrillar prote	C 479	7	1.0	243	2	JH0791	hypothetical prote
C 407	7	1.0	183	2	PN0109	keratin-like prote	C 480	7	1.0	244	2	S18948	centriocyclin precu
C 408	7	1.0	184	2	A42749	hypothetical prote	C 481	7	1.0	245	2	G72255	ABC transporter, A
C 409	7	1.0	185	2	S39317	replication initia	C 482	7	1.0	245	2	UC5346	cdd2 protein - Clo
C 410	7	1.0	185	2	S39318	replication initia	C 483	7	1.0	247	1	KYHUCM	chymase (EC 3.4.21
C 411	7	1.0	185	2	I40478	conserved hypotet	C 484	7	1.0	247	2	T07213	H+-transporting tw
C 412	7	1.0	187	2	B69536	acetylpolymyine am	C 485	7	1.0	247	2	S58394	myelin/oligodendro
C 413	7	1.0	188	2	C84787	hypothetical prote	C 486	7	1.0	247	2	H81719	conserved hypotet
C 414	7	1.0	188	2	F84629	hypothetical prote	C 487	7	1.0	247	2	T30008	hypothetical prote
C 415	7	1.0	189	2	H70416	antimicrobial prote	C 488	7	1.0	248	2	S33755	granzyme-like prot
C 416	7	1.0	190	2	S68230	hypothetical prote	C 489	7	1.0	249	2	B87309	hypothetical prote
C 417	7	1.0	191	2	T35637	probable ATP/GTP-b	C 490	7	1.0	250	2	A49910	extra protein - She
C 418	7	1.0	192	2	TN0577	D-aminobenzoic aci	C 491	7	1.0	250	2	AC1171	hypothetical prote
C 419	7	1.0	192	2	A70331	peptidyl-tRNA hydr	C 492	7	1.0	250	2	H83213	hypothetical prote
C 420	7	1.0	194	2	T47200	H+-transporting tw	C 493	7	1.0	251	2	S40995	hypothetical prote
C 421	7	1.0	194	2	T10939	auxin-induced prot	C 494	7	1.0	251	2	G70840	hypothetical prote
C 422	7	1.0	196	2	S72716	4-coumarate-CoA li	C 495	7	1.0	253	1	AB4582	type IV prelipin p
C 423	7	1.0	202	2	D86034	probable S-transfe	C 496	7	1.0	253	2	AB1717	dehydrogenase homo
C 424	7	1.0	202	2	E91187	probable S-transfe	C 497	7	1.0	253	2	AG1346	dehydrogenase homo
C 425	7	1.0	202	2	S47813	hypothetical prote	C 498	7	1.0	253	2	D83273	leader peptidase T
C 426	7	1.0	204	2	D91120	probable oxidoredu	C 499	7	1.0	253	2	T02106	hypothetical prote
C 427	7	1.0	204	2	D85965	probable oxidoredu	C 500	7	1.0	254	2	AH3218	hypothetical prote
C 428	7	1.0	205	2	S37804	hypothetical prote	C 501	7	1.0	254	2	T41954	hypothetical prote
C 429	7	1.0	205	2	D72781	hypothetical prote	C 502	7	1.0	256	2	S50280	PM21 protein - Ye
C 430	7	1.0	206	2	A53412	probable plasticqui	C 503	7	1.0	256	2	S50281	PM22 protein - Ye
C 431	7	1.0	206	2	S18250	collagen alpha 1(I	C 504	7	1.0	256	2	A90131	hypothetical prote
C 432	7	1.0	207	2	JH0145	hypothetical 24.1K	C 505	7	1.0	257	2	F87209	probable enoyl-CoA
C 433	7	1.0	209	2	B97543	hypothetical prote	C 506	7	1.0	257	2	C84890	hypothetical prote
C 434	7	1.0	209	2	AC2762	conserved hypotet	C 507	7	1.0	258	1	A64451	probable 3-isoprop
C 435	7	1.0	209	2	H65092	hypothetical prote	C 508	7	1.0	259	2	I53171	pancophylin - huma
C 436	7	1.0	210	2	S57943	cutinase - Phytroph	C 509	7	1.0	259	2	A71537	probable n-acetylm
C 437	7	1.0	210	2	S64376	hypothetical prote	C 510	7	1.0	259	2	F86475	hypothetical prote
C 438	7	1.0	214	2	T32066	hypothetical prote	C 511	7	1.0	260	2	B71234	probable transport
C 439	7	1.0	215	2	C69445	conserved hypotet	C 512	7	1.0	260	2	G75200	cobalt abc transpo
C 440	7	1.0	216	2	A71445	hypothetical prote	C 513	7	1.0	262	2	B81865	probable fibrinai
C 441	7	1.0	216	2	T45430	hypothetical prote	C 514	7	1.0	262	2	T34691	hypothetical prote
C 442	7	1.0	216	2	T46629	lp6 protein - lobl	C 515	7	1.0	262	2	C88325	protein F43G6.5 (I
C 443	7	1.0	217	2	T18603	hypothetical prote	C 516	7	1.0	263	2	D75593	conserved hypotet
C 444	7	1.0	218	2	T28689	hypothetical prote	C 517	7	1.0	264	2	T05450	hypothetical prote
C 445	7	1.0	221	2	A36370	immediate-early pr	C 518	7	1.0	266	2	S51833	arcelin-4 precuroso
C 446	7	1.0	221	2	T04592	glycine-rich celli	C 519	7	1.0	267	2	A57490	matrilysin (EC 3.4
C 447	7	1.0	222	2	T43500	hypothetical prote	C 520	7	1.0	268	2	S09860	hypothetical prote
C 448	7	1.0	223	2	C70524	hypothetical prote	C 521	7	1.0	269	2	C71228	hypothetical prote
C 449	7	1.0	223	2	AB1110	hypothetical prote	C 522	7	1.0	269	2	G86321	F644.11 protein -
C 450	7	1.0	224	2	T33691	hypothetical prote	C 523	7	1.0	270	2	A60830	Keratin, 70K type
C 451	7	1.0	225	2	T44720	hypothetical prote	C 524	7	1.0	270	2	D70171	ABC transporter, A
C 452	7	1.0	228	1	C46335	vif protein - Maed	C 525	7	1.0	270	2	E98296	hypothetical prote

C 526	7	1.0	270	2	AC2987	C 599	7	1.0	307	2	T39986	hypothetical prote
C 527	7	1.0	273	2	T08645	600	7	1.0	308	2	A83914	ABC transporter (p
C 528	7	1.0	273	2	F91083	601	7	1.0	309	2	T15696	hypothetical prote
C 529	7	1.0	273	2	H71212	602	7	1.0	310	2	H81161	cysteine synthase
C 530	7	1.0	274	2	C47700	603	7	1.0	310	2	T03200	probable RAV-like
C 531	7	1.0	274	2	H98174	604	7	1.0	313	2	S07924	alpha/beta-gliadin
C 532	7	1.0	274	2	AE3112	605	7	1.0	313	2	JE0328	intelectin - mouse
C 533	7	1.0	275	2	AC0069	606	7	1.0	313	2	H71438	hypothetical prote
C 534	7	1.0	275	2	D83626	607	7	1.0	314	2	C81735	tRNA delta-2-isope
C 535	7	1.0	276	2	F95259	608	7	1.0	314	2	G97379	phoe protein U5922
C 536	7	1.0	276	2	A99525	609	7	1.0	315	2	D48560	immediate-early pr
C 537	7	1.0	277	2	D86205	610	7	1.0	316	1	A38743	loricrin - human
C 538	7	1.0	278	2	H86923	611	7	1.0	316	2	T41365	hypothetical prote
C 539	7	1.0	278	2	D64750	612	7	1.0	316	2	E96770	hypothetical prote
C 540	7	1.0	278	2	T24430	613	7	1.0	319	2	A22364	alpha/beta-gliadin
C 541	7	1.0	279	2	T03830	614	7	1.0	320	2	T01848	hypothetical prote
C 542	7	1.0	280	2	H71320	615	7	1.0	320	2	T35265	probable D-amino a
C 543	7	1.0	281	2	T22670	616	7	1.0	320	2	C96922	probable acetyltra
C 544	7	1.0	281	2	T49537	617	7	1.0	321	2	T34065	hypothetical prote
C 545	7	1.0	282	2	H87251	618	7	1.0	321	2	C90402	transposase ISC117
C 546	7	1.0	286	1	EEWTA	619	7	1.0	321	2	D90326	transposase ISC117
C 547	7	1.0	286	1	S07923	620	7	1.0	324	2	B69521	hypothetical prote
C 548	7	1.0	286	2	F43570	621	7	1.0	324	2	I50125	cdx1 protein - com
C 549	7	1.0	288	2	T06282	622	7	1.0	325	2	T14227	NADH2 dehydrogenas
C 550	7	1.0	290	2	T31470	623	7	1.0	325	2	AF2665	conserved hypotet
C 551	7	1.0	290	2	T31201	624	7	1.0	326	2	D22364	alpha/beta-gliadin
C 552	7	1.0	290	2	T42526	625	7	1.0	326	2	E71047	probable signal re
C 553	7	1.0	291	2	AB1163	626	7	1.0	326	2	C90272	hypothetical prote
C 554	7	1.0	291	2	AB1522	627	7	1.0	327	2	G70789	probable epoxide h
C 555	7	1.0	292	2	B22364	628	7	1.0	328	2	G88499	protein K04G7.10 f
C 556	7	1.0	292	2	C71448	629	7	1.0	328	2	T36574	probable partition
C 557	7	1.0	293	2	H86718	630	7	1.0	328	2	H97089	uncharacterized AR
C 558	7	1.0	294	2	S23545	631	7	1.0	328	2	AF2597	hypothetical prote
C 559	7	1.0	294	2	A97937	632	7	1.0	329	2	E69456	hypothetical prote
C 560	7	1.0	294	2	A95069	633	7	1.0	330	2	B46835	kinasin-like prote
C 561	7	1.0	295	2	F59102	634	7	1.0	331	2	C95998	probable sugar upt
C 562	7	1.0	295	2	T51350	635	7	1.0	331	2	G75099	sugar transferase
C 563	7	1.0	296	2	A27319	636	7	1.0	332	2	T01483	hypothetical prote
C 564	7	1.0	296	2	S07361	637	7	1.0	333	2	E97447	hypothetical prote
C 565	7	1.0	296	2	S39587	638	7	1.0	333	2	T33649	hypothetical prote
C 566	7	1.0	296	2	G64877	639	7	1.0	333	2	F83107	hypothetical prote
C 567	7	1.0	296	2	E90862	640	7	1.0	334	2	B72033	hypothetical prote
C 568	7	1.0	296	2	D85756	641	7	1.0	334	2	C86593	conserved hypotet
C 569	7	1.0	296	2	AF0287	642	7	1.0	334	2	T36485	CT566 hypothetical
C 570	7	1.0	296	2	A10656	643	7	1.0	335	2	F72404	probable arac-fam1
C 571	7	1.0	296	2	T39265	644	7	1.0	336	2	S35156	flagellar motor sw
C 572	7	1.0	297	2	T06500	645	7	1.0	336	2	S72858	beta-glucanase - b
C 573	7	1.0	297	2	H83888	646	7	1.0	337	2	I47019	hypothetical prote
C 574	7	1.0	298	2	D95199	647	7	1.0	337	2	F47301	foliistatin - shee
C 575	7	1.0	298	2	B98066	648	7	1.0	339	2	A73473	Virb11 homolog - B
C 576	7	1.0	298	2	T51746	649	7	1.0	339	2	T25315	probable rRNA pyro
C 577	7	1.0	299	2	T00837	650	7	1.0	339	2	S55369	hypothetical prote
C 578	7	1.0	299	2	E82116	651	7	1.0	343	2	S45321	foliistatin - chic
C 579	7	1.0	299	2	T47989	652	7	1.0	343	2	T33945	hypothetical prote
C 580	7	1.0	300	2	F90274	653	7	1.0	344	1	A27701	foliistatin precus
C 581	7	1.0	300	2	T08453	654	7	1.0	344	2	T12337	NADH2 dehydrogenas
C 582	7	1.0	301	2	H97505	655	7	1.0	344	2	A33141	foliistatin 1 prec
C 583	7	1.0	301	2	AD2724	656	7	1.0	344	2	T45894	foliistatin - bov1
C 584	7	1.0	301	2	S72642	657	7	1.0	344	2	I57698	foliistatin - rat
C 585	7	1.0	301	2	G85928	658	7	1.0	345	2	T12339	NADH2 dehydrogenas
C 586	7	1.0	301	2	S35526	659	7	1.0	345	2	D88485	proetin F23F12.4 f
C 587	7	1.0	303	2	F89849	660	7	1.0	346	2	A12229	cobalam synthesis
C 588	7	1.0	303	2	F90301	661	7	1.0	346	2	T07660	probable serine/ch
C 589	7	1.0	304	2	A45989	662	7	1.0	348	2	G02297	gene N33 protein -
C 590	7	1.0	305	2	I52362	663	7	1.0	348	2	A83250	O-antigen chain 1e
C 591	7	1.0	305	2	A41812	664	7	1.0	348	2	S34494	ccsa protein - Eug
C 592	7	1.0	305	2	AC1588	665	7	1.0	348	2	T17304	hypothetical prote
C 593	7	1.0	305	2	AH1651	666	7	1.0	349	2	F84246	hypothetical prote
C 594	7	1.0	305	2	A596832	667	7	1.0	350	2	E87714	hypothetical prote
C 595	7	1.0	306	2	A54862	668	7	1.0	351	2	C75479	conserved hypotet
C 596	7	1.0	306	2	A72610	669	7	1.0	351	2	G83799	two-component sens
C 597	7	1.0	307	2	T51015	670	7	1.0	352	2	T06747	hypothetical prote
C 598	7	1.0	307	2	T03502	671	7	1.0	352	2	T35891	hypothetical prote

672	7	1.0	352	2	T33664	hypothetical prote
c 673	7	1.0	353	2	AD3169	transcription regu
674	7	1.0	353	2	H75020	hypothetical prote
c 675	7	1.0	355	2	T04317	homeobox protein L
c 676	7	1.0	355	2	S06939	hypothetical prote
c 677	7	1.0	356	2	T48206	protein kinase A1N
c 678	7	1.0	356	2	S39889	SWRI protein - Pod
c 679	7	1.0	356	2	H90168	GMP-binding protei
c 680	7	1.0	358	2	T26281	hypothetical prote
c 681	7	1.0	361	2	A48373	high-alkaline seri
c 682	7	1.0	361	2	G83756	subtilisin-type al
c 683	7	1.0	361	2	F91207	hypothetical prote
684	7	1.0	361	2	C72751	hypothetical prote
c 685	7	1.0	361	2	H86053	hypothetical prote
c 686	7	1.0	363	2	C86214	hypothetical prote
c 687	7	1.0	363	2	H92575	probable permease
688	7	1.0	363	2	AH2796	hypothetical prote
c 689	7	1.0	363	2	H81334	hypothetical prote
c 690	7	1.0	363	2	T51341	RNA helicase RH9 I
c 691	7	1.0	363	2	H97176	lysophospholipase
c 692	7	1.0	364	2	AH1857	anthranilate phosp
c 693	7	1.0	365	2	H96673	hypothetical prote
c 694	7	1.0	366	1	UC1106	inhibin alpha chai
c 695	7	1.0	366	1	A40056	inhibin alpha chai
c 696	7	1.0	366	2	F90474	hypothetical prote
c 697	7	1.0	367	2	E84453	probable GDSL-mot
c 698	7	1.0	368	2	G82321	twitching motility
699	7	1.0	369	2	A39157	probable RNA helic
700	7	1.0	369	2	S76435	hypothetical prote
701	7	1.0	369	2	B70968	hypothetical prote
c 702	7	1.0	371	2	T13021	hypothetical prote
c 703	7	1.0	371	2	H70326	hypothetical prote
c 704	7	1.0	372	2	T10202	hypothetical prote
c 705	7	1.0	373	1	DERTH1	3beta-hydroxy-Delt
c 706	7	1.0	373	1	DERTH2	3beta-hydroxy-Delt
c 707	7	1.0	373	1	DERTHM	3beta-hydroxy-Delt
c 708	7	1.0	373	2	A48769	3beta-hydroxyster
c 709	7	1.0	373	2	A37404	3beta-hydroxy-Delt
c 710	7	1.0	373	2	A49573	3beta-hydroxy-Delt
c 711	7	1.0	373	2	I49762	3beta-hydroxy-Delt
c 712	7	1.0	373	2	UC7289	G-protein coupled
c 713	7	1.0	374	2	E95361	probable muconate
c 714	7	1.0	375	2	AC0959	hypothetical prote
c 715	7	1.0	376	2	D84577	probable reca prot
c 716	7	1.0	378	2	T00481	probable RING zinc
c 717	7	1.0	379	1	QOEC4W	Citrate utilisatio
c 718	7	1.0	379	2	AP2409	mannosyl transfer
c 719	7	1.0	380	1	NPBY	hypothetical prote
c 720	7	1.0	380	2	G75170	anthranilate phosp
c 721	7	1.0	380	2	I50701	transcription fact
c 722	7	1.0	381	2	G96804	hypothetical prote
c 723	7	1.0	381	2	B86175	protein F19P13.21
724	7	1.0	381	2	A10622	hypothetical proph
725	7	1.0	382	2	T27058	hypothetical prote
c 726	7	1.0	383	2	A42111	Na+/H+-exchangin
c 727	7	1.0	386	2	F90112	DEAD box protein I
728	7	1.0	387	2	T48731	probable translati
729	7	1.0	389	2	D84068	RNA helicase BH334
c 730	7	1.0	389	2	A82327	MSHA biogenesis pr
731	7	1.0	390	2	G01936	Abi binding protei
c 732	7	1.0	391	2	G82330	probable oxygen-in
c 733	7	1.0	393	2	G96781	unknown protein F2
c 734	7	1.0	393	2	JC5275	voltage-gated pota
c 735	7	1.0	393	2	JC5614	RNA6 protein - rat
c 736	7	1.0	395	2	AC3571	acyl-CoA dehydroge
737	7	1.0	396	2	I38430	connexin45 - human
c 738	7	1.0	396	2	T18854	hypothetical prote
739	7	1.0	397	1	JC4017	triacylglycerol 11
c 740	7	1.0	397	2	B98293	mannanase dehydrat
c 741	7	1.0	397	2	AG2980	mannanase dehydrat
c 742	7	1.0	397	2	T17008	knotted-like home
c 743	7	1.0	397	2	T13516	hypothetical prote
c 744	7	1.0	398	2	A81717	conserved hypotet
745	7	1.0	398	2	T17009	knotted-like home
c 746	7	1.0	399	2	T32773	hypothetical prote
747	7	1.0	399	2	S54644	translation initia
c 748	7	1.0	399	2	C75024	hypothetical prote
c 749	7	1.0	399	2	B71140	hypothetical prote
c 750	7	1.0	399	2	T46267	probable zinc fing
c 751	7	1.0	401	2	A13568	mannanase dehydrat
c 752	7	1.0	403	2	A26371	sex steroid-bindin
c 753	7	1.0	403	2	A81464	translation initia
c 754	7	1.0	404	2	S54729	RNA-binding protei
c 755	7	1.0	404	2	G70886	probable integral
756	7	1.0	405	2	D84950	acetate kinase (EC
757	7	1.0	406	2	G97185	membrane protein o
c 758	7	1.0	407	2	AC2111	pilin biogenesis p
759	7	1.0	407	2	D82377	Para family protei
c 760	7	1.0	410	1	S68153	cellulase (EC 3.2.
c 761	7	1.0	410	2	S18157	globulin 2 - easer
c 762	7	1.0	411	2	S45142	translation initia
763	7	1.0	412	2	JC1453	translation initia
c 764	7	1.0	412	2	JC1452	translation initia
c 765	7	1.0	412	2	S62538	hypothetical coile
c 766	7	1.0	413	2	T47272	transposase tmpr I
767	7	1.0	413	2	T24767	hypothetical prote
c 768	7	1.0	413	2	AB3174	aminotransferase,
c 769	7	1.0	414	2	B96752	hypothetical prote
c 770	7	1.0	414	2	T38742	hypothetical prote
c 771	7	1.0	414	2	A40350	transcription repr
c 772	7	1.0	415	2	E83377	probable alcohol d
c 773	7	1.0	416	2	F83197	probable porin PA3
c 774	7	1.0	416	2	S50342	homeotic protein S
c 775	7	1.0	416	2	G83656	hypothetical prote
c 776	7	1.0	419	2	T43572	type III secretion
c 777	7	1.0	420	2	I51667	thrombin receptor
c 778	7	1.0	423	2	A90450	conserved hypotet
c 779	7	1.0	423	2	T03269	uroporphyrin-III C
780	7	1.0	425	2	S17759	protein kinase, ca
781	7	1.0	427	2	B64909	probable membrane
782	7	1.0	427	2	H90897	probable transport
783	7	1.0	427	2	G85719	probable transport
c 784	7	1.0	427	2	A32372	female-specific do
c 785	7	1.0	431	2	D86318	protein F15H16.6 I
c 786	7	1.0	431	2	F85097	hypothetical prote
c 787	7	1.0	431	2	T04868	hypothetical prote
c 788	7	1.0	432	2	T39816	hypothetical prote
c 789	7	1.0	432	2	D87649	colb protein limpo
c 790	7	1.0	432	2	T23561	hypothetical prote
c 791	7	1.0	433	1	QO8EV2	UL4 protein - hum
c 792	7	1.0	433	2	T11162	NADH2 dehydrogen
c 793	7	1.0	433	2	S07586	phosphopyruvate hy
c 794	7	1.0	433	2	S20963	homeotic protein H
c 795	7	1.0	434	2	S10246	phosphopyruvate hy
796	7	1.0	434	2	A23126	phosphopyruvate hy
c 797	7	1.0	434	2	S62168	hypothetical prote
c 798	7	1.0	435	2	AB1786	cellobiose phospho
799	7	1.0	435	2	AC1410	hypothetical prote
c 800	7	1.0	435	2	S75346	hypothetical prote
c 801	7	1.0	435	2	T08454	hypothetical prote
c 802	7	1.0	436	2	JC5599	cholecystokinin-A
c 803	7	1.0	437	2	S59151	NADH2 dehydrogen
c 804	7	1.0	438	2	AD0389	phosphate regulon
c 805	7	1.0	438	2	T07790	transaldolase (EC
c 806	7	1.0	439	2	B81410	probable integral
c 807	7	1.0	439	2	E84780	hypothetical prote
c 808	7	1.0	440	2	T11358	NADH2 dehydrogen
809	7	1.0	441	2	T51347	RNA helicase RH23
810	7	1.0	441	2	D71191	hypothetical prote
811	7	1.0	442	2	F75150	ubp-interacting pr
812	7	1.0	442	2	A71433	hypothetical prote
813	7	1.0	442	2	H98243	agae protein limpo
814	7	1.0	442	2	AD3042	oxidoreductase Atu
c 815	7	1.0	443	2	S37612	NADH2 dehydrogen
c 816	7	1.0	444	2	T42674	hypothetical prote
817	7	1.0	445	2	S39203	phosphopyruvate hy

C 818	7	1.0	445	2	A57400	embryonic TEA doma
C 819	7	1.0	445	2	T05639	hypothetical prote
C 820	7	1.0	446	2	T45525	WSC4 homolog (impo
C 821	7	1.0	447	2	AF2145	hypothetical prote
C 822	7	1.0	448	2	E89936	hypothetical prote
C 823	7	1.0	449	1	H70526	probable cytochrom
C 824	7	1.0	449	2	I64246	ATP-dependent RNA
C 825	7	1.0	450	2	S73545	ATP-dependent RNA
C 826	7	1.0	450	2	H87493	conserved hypotet
C 827	7	1.0	451	2	S56717	calcium-dependent
C 828	7	1.0	451	2	G82096	sodium-dependent t
C 829	7	1.0	452	2	C71391	NADH2 dehydrogenas
C 830	7	1.0	452	2	B87075	probable DNA-bindi
C 831	7	1.0	452	2	D70768	hypothetical prote
C 832	7	1.0	456	2	B90543	hypothetical prote
C 833	7	1.0	459	2	T11110	NADH2 dehydrogenas
C 834	7	1.0	459	2	D96833	hypothetical prote
C 835	7	1.0	459	2	AE0679	probable membrane
C 836	7	1.0	459	2	B95009	potassium uptake p
C 837	7	1.0	459	2	F97880	hypothetical prote
C 838	7	1.0	460	2	T13881	NADH2 dehydrogenas
C 839	7	1.0	461	2	H84099	cell wall-binding
C 840	7	1.0	463	2	TS1194	hypothetical prote
C 841	7	1.0	464	2	JH0824	gamma-aminobutylic
C 842	7	1.0	464	2	S22697	extensin - Volvox
C 843	7	1.0	465	1	T03024	calcium-dependent
C 844	7	1.0	465	1	WZBE84	51.3K capsid prote
C 845	7	1.0	465	2	AC0347	probable membrane
C 846	7	1.0	465	2	T34984	probable integral
C 847	7	1.0	465	2	T19113	hypothetical prote
C 848	7	1.0	465	2	T27415	hypothetical prote
C 849	7	1.0	466	2	C42360	cellulase (EC 3.2.
C 850	7	1.0	467	2	S38535	cytochrome P450 76
C 851	7	1.0	469	1	I38950	DNA-directed DNA p
C 852	7	1.0	469	2	E82264	probable polysacch
C 853	7	1.0	471	2	T20690	hypothetical prote
C 854	7	1.0	473	2	E72723	probable RNA-bind
C 855	7	1.0	473	2	G82405	transcription regu
C 856	7	1.0	474	1	T05239	acid phosphatase (
C 857	7	1.0	474	2	G84372	dihydroliipoamide d
C 858	7	1.0	474	2	S63658	NADH2 dehydrogenas
C 859	7	1.0	474	2	T39587	serine/threonine-p
C 860	7	1.0	475	1	A56824	dihydroliipoamide d
C 861	7	1.0	476	2	C96577	hypothetical prote
C 862	7	1.0	476	2	H84524	probable fatty aci
C 863	7	1.0	476	2	T32032	hypothetical prote
C 864	7	1.0	478	2	G75530	probable ATP-depen
C 865	7	1.0	480	2	F86207	hypothetical prote
C 866	7	1.0	480	2	A31589	carboxypeptidase C
C 867	7	1.0	481	2	JE0377	p70 S6 kinase (EC
C 868	7	1.0	481	2	A35628	loricrin - mouse
C 869	7	1.0	481	2	C71079	hypothetical prote
C 870	7	1.0	482	2	S60757	catalase (EC 1.11.
C 871	7	1.0	482	2	S27608	cysteine proteinas
C 872	7	1.0	483	2	H87492	yjer family protei
C 873	7	1.0	484	2	T05650	calcium-dependent
C 874	7	1.0	484	2	F88924	protein R03C2.2 [i
C 875	7	1.0	485	2	Q01957	glucagon receptor
C 876	7	1.0	485	2	AG2531	hypothetical prote
C 877	7	1.0	487	1	S71770	calcium-dependent
C 878	7	1.0	487	1	E69895	xylinokinas homol
C 879	7	1.0	487	2	AB1204	glycerol kinase ho
C 880	7	1.0	488	2	S18156	gluculin 1 - easte
C 881	7	1.0	489	1	NIBCAT	nitrogenase (EC 1.
C 882	7	1.0	489	1	AG3038	conserved hypotet
C 883	7	1.0	490	1	S71776	calcium-dependent
C 884	7	1.0	490	2	D70008	nicotinate phospho
C 885	7	1.0	490	2	JC5641	sugar transporter
C 886	7	1.0	490	2	F89824	hypothetical prote
C 887	7	1.0	491	2	JC6197	stromelysin 3 (EC
C 888	7	1.0	491	2	T48991	hypothetical prote
C 889	7	1.0	491	2	S63206	LE11 protein homol
C 890	7	1.0	492	1	T03271	calcium-dependent
C 891	7	1.0	492	2	A44399	stromelysin 3 (EC
C 892	7	1.0	493	2	S78183	NADH2 dehydrogenas
C 893	7	1.0	493	2	G84263	long-chain fatty-a
C 894	7	1.0	494	2	AB1326	two-component resp
C 895	7	1.0	496	2	JC5110	cyclin-dependent k
C 896	7	1.0	497	1	A43349	steroid 21-monoxy
C 897	7	1.0	497	1	S53834	NADH2 dehydrogenas
C 898	7	1.0	497	1	T14433	reproductive meris
C 899	7	1.0	497	2	S22708	homeotic protein e
C 900	7	1.0	498	2	T11039	NADH2 dehydrogenas
C 901	7	1.0	498	2	T13434	hypothetical prote
C 902	7	1.0	500	2	A53658	hypothetical prote
C 903	7	1.0	500	2	AC1915	prostaglandin-I by
C 904	7	1.0	502	2	AD0691	NADH dehydrogenase
C 905	7	1.0	504	2	T07415	conserved hypotet
C 906	7	1.0	506	1	W2ML47	probable serine/th
C 907	7	1.0	507	2	S75961	E2 protein - human
C 908	7	1.0	507	2	T31975	NADH2 dehydrogenas
C 909	7	1.0	507	2	T24944	hypothetical prote
C 910	7	1.0	509	2	T39542	hypothetical prote
C 911	7	1.0	511	2	T11940	ribosomal protein
C 912	7	1.0	512	2	T40576	probable structure
C 913	7	1.0	513	1	T02259	calcium-dependent
C 914	7	1.0	513	2	T46788	hypothetical prote
C 915	7	1.0	514	2	T10938	calcium-dependent
C 916	7	1.0	514	2	T44976	tydantoinsase homol
C 917	7	1.0	515	2	D98247	next protein (AF11
C 918	7	1.0	515	2	AC2635	glycosyltransferas
C 919	7	1.0	516	2	AC1540	ATP-dependent RNA
C 920	7	1.0	516	2	S40996	hypothetical prote
C 921	7	1.0	518	2	T11456	cytochrome-c oxida
C 922	7	1.0	518	2	T25507	hypothetical prote
C 923	7	1.0	518	2	S46784	hypothetical prote
C 924	7	1.0	519	2	I54523	kinesin-related pr
C 925	7	1.0	520	2	I51556	recombination acti
C 926	7	1.0	520	2	AB1183	ATP-dependent RNA
C 927	7	1.0	521	1	HYBS	bacillolysin (EC 3
C 928	7	1.0	521	2	G96543	calcium-dependent
C 929	7	1.0	521	2	G02129	bacillolysin (EC 3
C 930	7	1.0	521	2	T01242	peptide transport
C 931	7	1.0	522	2	D82195	phage replication
C 932	7	1.0	523	2	D90270	maltoase ABC transp
C 933	7	1.0	525	2	T31802	hypothetical prote
C 934	7	1.0	525	2	T48824	hypothetical prote
C 935	7	1.0	526	1	KRHOVI	keratin, 54K type
C 936	7	1.0	526	2	E65024	Hydrogenase-4 comp
C 937	7	1.0	526	2	D91047	hydrogenase 4 memb
C 938	7	1.0	526	2	H85891	hydrogenase 4 memb
C 939	7	1.0	526	2	C83790	hypothetical prote
C 940	7	1.0	527	2	S26037	NADH2 dehydrogenas
C 941	7	1.0	528	2	G02127	fus-like protein -
C 942	7	1.0	528	2	S42510	Rag-2 protein - ch
C 943	7	1.0	529	2	S46116	probable regulator
C 944	7	1.0	530	2	T48627	hypothetical prote
C 945	7	1.0	531	1	T02993	calcium-dependent
C 946	7	1.0	531	2	D85059	probable calcium d
C 947	7	1.0	531	2	B83082	probable binding p
C 948	7	1.0	532	2	T14335	protein kinase, ca
C 949	7	1.0	533	1	S56652	calcium-dependent
C 950	7	1.0	533	1	JS0304	developmental cont
C 951	7	1.0	534	1	JC1515	calcium-dependent
C 952	7	1.0	534	2	H71069	probable nodulatio
C 953	7	1.0	538	2	T06874	calcium-dependent
C 954	7	1.0	538	2	S67766	RNA-export mediat
C 955	7	1.0	539	2	B72707	hypothetical prote
C 956	7	1.0	540	1	T01989	calcium-dependent
C 957	7	1.0	541	1	S51799	nucleoporin NUP57
C 958	7	1.0	542	1	S56651	calcium-dependent
C 959	7	1.0	542	2	AG0341	hypothetical prote
C 960	7	1.0	542	2	D83041	probable two-compo
C 961	7	1.0	544	2	D84550	probable calmoduli
C 962	7	1.0	544	2	S43058	Cteta protein eta
C 963	7	1.0	545	2	T45935	probable galactose

964	7	1.0	546	2	A84630	hypothetical prote
965	7	1.0	547	2	B97114	probable kinase re
966	7	1.0	547	2	S44841	KO6H7.1 protein -
967	7	1.0	548	2	T47510	probable transport
968	7	1.0	549	2	B32372	male-specific doub
969	7	1.0	550	2	G70597	probable proteinase
970	7	1.0	551	2	S65289	hypothetical prote
971	7	1.0	551	2	S66740	probable transcrip
972	7	1.0	553	1	T02139	calcium-dependent
973	7	1.0	553	1	B55483	transcription init
974	7	1.0	554	1	T03263	calcium-dependent
975	7	1.0	554	2	T05476	calcium-dependent
976	7	1.0	554	2	T22140	hypothetical prote
977	7	1.0	554	2	E87375	conserved hypocher
978	7	1.0	554	2	T50118	kinesin-related pr
979	7	1.0	556	2	T06126	calcium-dependent
980	7	1.0	556	2	G90352	thermopsin precurs
981	7	1.0	556	2	T16790	hypothetical prote
982	7	1.0	557	1	NUVKL	glucose-6-phosphat
983	7	1.0	559	2	T22179	hypothetical prote
984	7	1.0	559	2	JT0949	egg-specific prote
985	7	1.0	560	2	F81423	L-lactate permease
986	7	1.0	560	2	T08433	helicase homolog h
987	7	1.0	562	2	C71604	RNA helicase PRB08
988	7	1.0	563	2	T51348	AS5R protein - vac
989	7	1.0	564	2	C42523	Salp17R protein -
990	7	1.0	564	2	J01792	lyrs protein slr10
991	7	1.0	565	2	S75255	protein F1504.19
992	7	1.0	565	2	H86477	lestin specific ba
993	7	1.0	570	2	G02753	PPR-repeat protein
994	7	1.0	570	2	C86440	hypothetical prote
995	7	1.0	572	2	T16780	conserved hypocher
996	7	1.0	572	2	H75271	probable homeodoma
997	7	1.0	575	2	T48224	cytochrome c-type
998	7	1.0	577	2	S77715	probable acyl-CoAs
999	7	1.0	578	2	A70877	probable Ompa-fam1
1000	7	1.0	578	2	AG0119	

## ALIGNMENTS

## RESULT 1

146407 probable RNA helicase protein DKFzp434B122.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 17-Nov-2000

C:Accession: T46407

R:Blum, H.; Baurerachs, S.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23034

A:Accession: T46407

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-635 <AAA>

A:Cross-references: EMBL:AL137462; PIDN:CA870750.1

A:Experimental source: adult testis; clone DKFzp434B122

C:Genetics:

A>Note: DKFzp434B122.1

## Alignment Scores:

Pred. No.:	0	Length:	635
Score:	522.00	Matches:	622
Percent Similarity:	99.84%	Conservative:	0
Best Local Similarity:	99.84%	Mismatches:	1
Query Match:	72.10%	Gaps:	0
DB:	2		

US-09-714-865-15 (1-2172) x T46407 (1-635)

QY 268 AGTTTGAAGACAGAGGTTTTCACACAGAGTTTGAAGATGATAGCTCTGTTTC 327

Db 1 SerpneGlyAsnAGrGlyPheSerAsnSerArpGheGlyAspGlyAspSerSerglyPhe 20

QY	328	TGAGAGAGCTCTAGTAATGACTCGAAGATAATCCAAACGGAACAGAGGTTTTCAG 387
Db	21	TPARGGlucSerSerAsnAspCysGluAspAsnProThrArgAsnArgGlyPheSerLys 40
QY	388	AGAGGGCGCTATCGAGATGGAATTAATTCAGAACTTCAGGGCCATACGAAAGAGTGA 447
Db	41	ArgGlyGlyTYrArgAspGlyAsnAsnSerGluAspGlyProTYrArgArgGlyGly 60
QY	448	AGAGGTAGTTTCCAGAGTCCCGTGGAGATTGGCTCGAAGATCCAAATATGACTTA 507
Db	61	ArgGlySerPheArgGlyCysArgGlyGlyPheGlyLeuGlySerProAsnAsnAspLeu 80
QY	508	GAACCAAGCAATGTATGACGCGACTGGTGGCTTTTGGTTCTGAAAGACCATATTA 567
Db	81	AspProAspGluCysSerGlnArgThrGlyGlyLeuPheGlySerArgArgProValLeu 100
QY	568	AGTGGCAGAGTATGATGATCTTTCACAAAGAGTGGCAGTGAAGTGAACAGAGT 627
Db	101	SerGlyTYrGlyAsnGlyAspThrSerGlnSerArgSerglySerglySerglyArgGly 120
QY	628	GGTTACAAAGGTTTAATGAGAGTAATACAGGCTCGAAGAAATCTTGAAGATCA 687
Db	121	GlyTYrLysGlyLeuAsnGluGluValIleThrGlySerglyLysAsnSerTrpLysSer 140
QY	688	GAGCAGAAAGAGGAGAAAGTATGATGATCTCAAGACCAAAAGTCACTACATACCCCT 747
Db	141	GluAlaGluGluGlyGlyGluSerSerAspThrGlnGlyProValThrTrpIleProPro 160
QY	748	CCTCCAGCTGAGAGATGAGATCTTCATCTTTGCAATTAACAGCAGCATTAACCTTGAC 807
Db	161	ProProProGluAspGluAspSerIlePheAlaIleTYrGlyThrGlyLysAsnPheAsp 180
QY	808	AAATACAGACATTTCTTGGAAGTCTGGACATGATGACACACAGCAATTCGACT 867
Db	181	LysTYrAspThrIleLeuValGluValSerglyLysAspAlaProValIleLeuThr 200
QY	868	TTTGAAGAGCTATCTCTGTCAGACATGAAATACACATTCCTAAGCTGGTTACT 927
Db	201	PheGluGluAlaAsnLeuAsnGlnThrLeuAsnAsnAlaAlaLysAlaGlyTYrThr 220
QY	928	AACTTACTCTCTGTCGAAAATACAGTATTCCTATATCTTCAGCAGAGAGATTTGATG 987
Db	221	LysLeuThrProValGlnLysTYrSeriIleProIleIleLeuAlaGlyArgAspLeuMet 240
QY	988	GCCTTGCTCAACAGAGTCTGGAGACATCGGGCTTTCTCCTACCAATTTGGCTCAT 1047
Db	241	AlaCysAlaGlnThrIleSerglyLysThrAlaAlaPheLeuLeuProIleLeuAlaHis 260
QY	1048	ATGATCATGATGGAATTAATGACAGTCTGTTTAAAGATTGACGAGACAGAGTGTATT 1107
Db	261	MetMetHisAspGlyIleThrIleAspArgPheLysGluLeuGlnIleProGluCysIle 280
QY	1108	ATTGTGACCAACCTGAGATTTGTCAACACAGATTATTTGGAGCCGAAATTTTCT 1167
Db	281	IleValAlaProThrArgGluLeuValIaangIleIleTYrLeuGluAlaArgLysPheSer 300
QY	1168	TTTGGGACTTGTAGAGCTGTGTTATATATTTGGGGAAACCAAGCTGGAGACTCAATT 1227
Db	301	PheGlyThrCysValArgAlaValValIleTYrGlyGlyThrGlnLeuGlyHisSerIle 320
QY	1228	CGACAATATAGTACAGAGCTGTATATATATATATGCTACTCCGGAAGACTGATGATATC 1287
Db	321	ArgGlnIleValGlnIleCysAsnIleLeuGlyAlaThrProGlyArgLeuMetAspIle 340
QY	1288	ATRGGCAAAAGAAAGATTGGTCTCAACAGATCAAAATCTTACTTTTGGATGAAGCTGAT 1347
Db	341	IleGlyLysGluLysIleGlyLeuLysGlnIleIleTYrLeuValLeuAspGluAlaAsp 360
QY	1348	CGCATGTGATATGGGTTTGGTCCAGAAAGAAAGTAAATTTCTTGCCAGGAATG 1407
Db	361	ArgMetLeuAspPheCylPheGlyProGluMetLysLeuLeuIleSerCysProGlyMet 380
QY	1408	CCATCAAGGAACAGCGCCAAACCTTATGTTTCAGTGAACCTTTTCAGAGAAATTCAA 1467



```

Db      381 ProSeLySgLnInrGInThrLeuMePheSeRAlaThrPheProGluGluIleGln 400
QY      1468 AGGTGGTGCAGAGATTGTTAAAGTCAAAATATCTGTTGTCGTGTGGACAAGGGG 1527
Db      401 ArgLeuAlaAlaGluPheLeuLysSerAsnTrpLeuPheValAlaValGlyGlnValGly 420
QY      1528 GGAGCATGTAGAGATGTCAGACGCGTTCCTCAAGTTGGCCAGTTCTCAAAAAGAA 1587
Db      421 GlyAlaCysArgAspValGlnGlnThrValLeuGlnValGlyGlnPheSerLysArgGlu 440
QY      1588 AGCTGTTGAATTCGTGCGAAACATAGGGGATGAAAGAACTATGGCTTTGTTGAAACT 1647
Db      441 LysLeuValGluIleLeuArgAsnIleGlyAspGluArgThrMetValPheValGluThr 460
QY      1648 AAGAAAAAGCAATTTTCTGCACTTTCTTTGCAAGAAAAATATCAATCAACAAG 1707
Db      461 LysLysLysAlaAspPheIleAlaThrPheLeuCysGlnGluLysIleSerThrTrpSer 480
QY      1708 ATCCATGTGATCGGGAACAGAGAGCGGAGCAAGCTCTTGGAGATTTCGCTTTGGA 1767
Db      481 IleHisGlyAspArgGluGlnArgGluArgGluGlnAlaLeuGlyAspPheArgPheGly 500
QY      1768 AAGTCCCAAGTTCTTGTGTCTACTAGTACGTGCGAGAGGCGTGAATTTGAAATGTG 1827
Db      501 LysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluAsnVal 520
QY      1828 CAACATGTTATCAATTTTATCTCTCTTCACATGATGAATATGTTATCAATGAG 1887
Db      521 GlnHisValIleLeuAsnPheAspLeuProSerThrIleAspIleuTrpValHisArgIleGly 540
QY      1888 CGTACTGTGTCGTTGTGGGAATACTGCGACAGACAAATTTCTTTTGTATCTTGAATCGAT 1947
Db      541 ArgThrGlyArgGlyGlyAsnThrGlyArgGlyAlaIleSerPhePheAspLeuGluSerAsp 560
QY      1948 AACCATTTAGCAGCGCTCTAGTAAAGTATTTGACAGATCTCAACAGATGTTCTGCA 2007
Db      561 AsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValProAla 580
QY      2008 TGGTTGGAGAAATTCCTTTATACATACATTCCTCGGCTTCAGTGTAGTAAAGAGA 2067
Db      581 TrpLeuGlnGlnIleAlaPheSerThrTrpIleProGlyPheSerGlySerThrAspGly 600
QY      2068 AAGCTTTTCATCAGTTGATACAGAAAGGAGGAGAGACTTTGAACAGCTGGGTT 2127
Db      601 AsnValPheAlaSerValAspTrpArgLysGlyLysSerThrLeuAsnThrAlaGlyPhe 620
QY      2128 TCTTCTTCA 2136
Db      621 SerSerSer 623

RESULT 2
JC2534
RVLG protein - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 20-May-1995 #sequence_revision 14-Jul-1995 #text_change 02-Feb-2001
C/Accession: JC2534
R/Komiyu, T.; Tanigawa, Y.
Biochem. Biophys. Res. Commun. 207, 405-410, 1995
A/Title: Cloning of a gene of the DEAD Box protein family which is specifically expressed
A/Reference number: JC2534, PMID:95160706; PMID:7857296
A/Accession: JC2534
A/Molecule type: mRNA
A/Residues: 1-713 <KOM>
A/Cross-references: GB:575375; NID:9806463; PIND:AA03364.1; PID:9806464
C/Comment: This protein contains a conserved DEAD box.
C/Keywords: ATP; nucleotide-binding; P-loop
F/317-324/Region: nucleotide-binding motif A (P-loop)
F/427-432/Region: nucleotide-binding motif B
F/431-434/Region: DEAD motif

Alignment Scores:
Pred. No.: 1,57e-74 Length: 713

```

```

Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x JC2534 (1-713)
QY      718 CAAGACCAAAAGTAGCTACATACCCCTCTCCACCTGAGATGAGAGACTCCATCTTT 777
Db      225 GlnGlyProLysValThrTrpIleProProProProProGluAspLysPheSerIlePhe 244
QY      778 GCACATTATCAGACAGGCAATAAATCTGCAAAATACGACATATTTCTGTGGAAGTGTCT 837
Db      245 AlaHisTrpGlnThrGlyIleAsnPheAspLysTrpAspThrIleLeuValGluValSer 264
QY      838 GCACATGATGCACACACAGCAATTTCTGACTTTTGAAGAAGCTAATCTCTGTGAGACACTG 897
Db      265 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuGlyGlnThrLeu 284
QY      898 AATAACCAATGCTTAAAGCTGTTATACATTAAGCTTACTCTGTGCAAAAATACAGATT 957
Db      285 AsnAsnAsnIleAlaLysAlaGlyTrpThrLysLeuThrProValGlnLysTrpSerIle 304
QY      958 CCTATC 963
Db      305 ProIle 306

RESULT 3
I49638
probable RNA helicase protein - mouse (fragment)
N/Alternate names: Drosophila vasa homolog
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Feb-2001
C/Accession: I49638
R/Fujisawa, Y.; Komiyu, T.; Kawabata, H.; Sato, M.; Fujimoto, H.; Furusawa, M.; Noce, T.
Proc. Natl. Acad. Sci. U.S.A. 91, 12258-12262, 1994
A/Title: Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosophila
A/Reference number: I49638, PMID:95083681; PMID:7991615
A/Accession: I49638
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-637 <RES>
A/Cross-references: GB:D14859; NID:9286074; PIND:BA03584.1; PID:9286075
C/Keywords: ATP; nucleotide-binding; P-loop
F/346-253/Region: nucleotide-binding motif A (P-loop)
F/356-361/Region: nucleotide-binding motif B
F/360-363/Region: DEAD motif

Alignment Scores:
Pred. No.: 3,76e-56 Length: 637
Score: 64.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.84% Indels: 0
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x I49638 (1-637)
QY      718 CAAGACCAAAAGTAGCTACATACCCCTCTCCACCTGAGATGAGAGACTCCATCTTT 777
Db      154 GlnGlyProLysValThrTrpIleProProProProProGluAspLysPheSerIlePhe 173
QY      778 GCACATTATCAGACAGGCAATAAATCTGCAAAATACGACATATTTCTGTGGAAGTGTCT 837
Db      174 AlaHisTrpGlnThrGlyIleAsnPheAspLysTrpAspThrIleLeuValGluValSer 193
QY      838 GCACATGATGCACACACAGCAATTTCTGACTTTTGAAGAAGCTAATCTCTGTGAGACACTG 897
Db      194 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuGlyGlnThrLeu 213
QY      898 AATAACCAAT 909

```

Db 214 Asmbanentile 217

RESULT 4

151235  
DEAD box protein - African clawed frog (fragment)  
C/Species: Xenopus laevis (African clawed frog)  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Feb-2001  
C/Accession: 151235  
R/Komiyama, T.; Itoh, K.; Ikenishi, K.; Furusawa, M.  
Dev. Biol. 162, 354-363, 1994  
A/Title: Isolation and characterization of a novel gene of the DEAD box protein family  
A/Reference number: 151235; MUID:94200507; PMID:8150200  
A/Accession: 151235  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-700 <KOM>  
A/Cross-references: GB:S69534; NID:g2896106; PIDN:AAC03114.1; PID:g2896107  
C/Genetics:  
A/Name: XVLG1  
C/Keywords: ATP; nucleotide binding; P-loop  
F/318-335/Region: nucleotide-binding motif A (P-loop)  
F/428-433/Region: nucleotide-binding motif B  
F/432-435/Region: DEAD motif

Alignment Scores:

Pred. No.:	2,51e-15	Length:	700
Score:	24.00	Matches:	24
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.31%	Indels:	0
DB:	2	Gaps:	0

US-09-714-865-15 (1-2172) x 151235 (1-700)

QY 970 GCAGGACGAGATTGATGCTTGCTCAACAGGGCTCGGAGAGACTGCGCTTTCTC 1029  
|||  
Db 310 AaagtyaagaaplemerAlaCyaalaglnhrtyserglyvtrAlaAlaPheleu 329  
|||  
QY 1030 CTACCAATTGTTG 1041  
|||  
Db 330 Leuprolleu 333  
|||

RESULT 5

A58768  
ATP-dependent RNA helicase homolog - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 31-Dec-1990 #sequence\_revision 17-Apr-1998 #text\_change 19-Jan-2001  
C/Accession: A58768; S01676; S10129; A31922  
R/Liang, L.; Diehl-Jones, W.; Lasko, P.  
unpublished results, 1995, cited by GenBank in release 103.0  
A/Description: Localization of Vasa protein to the Drosophila pole plasma is independent  
A/Reference number: A58768  
A/Accession: A58768  
A/Status: not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 1-661 <LAS1>  
A/Cross-references: EMBL:X12945; NID:g433675; PIDN:CAA31405.1; PID:g1054723  
R/Lasko, P.F.; Ashburner, M.  
Nature 335, 611-617, 1988  
A/Title: The product of the Drosophila gene vasa is very similar to eukaryotic initiator  
A/Reference number: S01676; MUID:89014721; PMID:3140040  
A/Accession: S01676  
A/Molecule type: DNA  
A/Residues: 1-34, 'R', 36-448, 'LRF', 453-460, 'R', 462-589, 'Q', 591-661 <LAS2>  
A/Cross-references: EMBL:X12945; NID:g433675  
R/Ashburner, M.  
submitted to the EMBL Data Library, October 1988  
A/Reference number: S10129  
A/Accession: S10129  
A/Molecule type: DNA  
A/Residues: 1-34, 'R', 36-644, 'Q', 645-661 <ASH>  
A/Cross-references: EMBL:X12945; NID:g433675  
F/Hay, B.; Jan, L.Y.; Jan, Y.N.

```

Cell 55, 577-587, 1988
A:Title: A protein component of Drosophila polar granules is encoded by vasa and has exte
A:Reference number: A19222; MUID:89028669; PMID:3052853
A:Accession: A19222
A:Molecule type: mRNA
A:Residues: 1-34, 'R', 36-154, 168-264, 'Y', 266-321, 'C', 323-451, 'F', 453-581, 'R', 583-593, 'H', 5
A:Cross-references: GB:J35160; NID:G158795; PIDN:AAA29013.1; PID:G158796
A>Note: the authors translated the codon TGT for residue 322 as Val
C:Genetics:
A:Gene: vasa
A:Cross-references: FlyBase:FBgn0003970
A:Introns: 8/3; 138/1; 234/3; 482/3; 554/1; 644/2
C:Superfamily: ATP-dependent RNA helicase DBP1
C:Keywords: ATP; nucleotide binding; P-loop
F:281-236/Region: nucleotide-binding motif A (P-loop)
F:395-400/Region: nucleotide-binding motif B
F:399-402/Region: DEAD motif

Alignment Scores:
Pred. No.: 2,65e-14 Length: 661
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3,18* Indels: 0
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x A58768 (1-661)

OY 973 GGACGAGTTGATGGCTTGCTCAACAGGGTCTGGAGAGCTCGGCTTTCTCTA 1032
|||||
Db 282 GtargapleuMeAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeu 301
OY 1033 CCAATTTTG 1041
|||||
Db 302 Prolleu 304

RESULT 6
A32378
Probable ATP-dependent RNA helicase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Sep-1989 #sequence_revision 08-Sep-1989 #ext_change 19-Jan-2001
C:Accession: A32378
R:eroy, P.; Alzari, P.; Sassoon, D.; Wolgemuth, D.; Fellous, M.
Cell 57, 549-559, 1989
A:Title: The protein encoded by a murine male germ cell-specific transcript is a putative
A:Reference number: A32378; MUID:89249320; PMID:2720782
A:Accession: A32378
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-660 <LER>
A:Cross-references: GB:J04847; NID:G200388; PIDN:AAA39942.1; PID:G200389
C:Superfamily: ATP-dependent RNA helicase DBP1
C:Keywords: ATP; DNA binding; nucleotide binding; P-loop
F:223-230/Region: nucleotide-binding motif A (P-loop)
F:342-347/Region: nucleotide-binding motif B
F:346-349/Region: DEAD motif

Alignment Scores:
Pred. No.: 2,78e-13 Length: 660
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3,04* Indels: 0
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x A32378 (1-660)

OY 976 CGAGATTGATGGCTTGCTCAACAGGGTCTGGAGAGCTCGGCTTTCTCTACCA 1035
|||||
Db 217 ArgapleuMeAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuPro 236
OY 1036 ATTTTG 1041
|||||

```

Db 237 I1e1eu 238

RESULT 7  
184741  
RNA helicase - mouse  
N:Alternate names: RNA helicase ERH  
C:Species: Mus musculus (house mouse)  
C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 19-Jan-2001  
C/Accession: 184741; S56112  
R:Gene: S.L.; Conboy, J.G.  
Gene 140, 171-177, 1994  
A>Title: Mouse erythroid cells express multiple putative RNA helicase genes exhibiting h  
A/Reference number: 149731; MUID:94192995; PMID:8144024  
A/Accession: 184741  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A/Residues: 1-662 <RES>  
A/Cross-references: GB:I25126; NID:9407995; PIDN:AAA53630.1; PID:9407996  
A/Genetics: RES1  
R:Snowden, J.; Putt, W.; Morrison, K.; Beddington, R.; Edwards, Y.  
Biochem. J. 308, 839-846, 1995  
A>Title: The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of R  
A/Reference number: S56112; MUID:97104282; PMID:8948440  
A/Accession: S56112  
A:Molecule type: mRNA  
A/Residues: 1-229, 'R', 231-662 <SOW>  
A/Cross-references: GB:238117; NID:91835121; PID:91835122  
A/Experimental source: strain c57bl/6; notochord  
A/Genetics: SOW1  
A/Note: the sequence is revised in GenBank entry MMBRNHL, release 117, (PID:1835122)  
C/Genetics: <RES>  
A/Note: the revised sequence is now identical to PIR accession 184741  
A:Gene: Bif4a-r1; MGI:Ddx19  
A/Cross-references: MGI:99526  
C/Genetics: <SOW>  
A:Gene: MGI:Ddx3  
A/Cross-references: MGI:103064  
A/Map position: 1  
C:Superfamily: ATP-dependent RNA helicase DBP1  
C/Keywords: ATP; nucleotide binding; P-loop  
F:224-231/Region: nucleotide-binding motif A (P-loop)  
F:343-348/Region: nucleotide-binding motif B  
F:347-350/Region: DEAD motif

Alignment Scores:

Pred. No.:	2,786-13	Length:	662
Score:	22.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.04%	Indels:	0
DB:	1	Gaps:	0

US-09-714-865-15 (1-2172) x 184741 (1-662)

Qy 976 CGAGATTGATGCTTGCTCAACAGGCTGGAGAGCTGGCTTTCTCTACCA 1035  
|||||  
Db 218 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuPro 237  
|||||

Qy 1036 ATTTTG 1041  
|||||  
Db 238 I1e1eu 239

RESULT 8  
S13654  
ATP-dependent RNA helicase - African clawed frog  
C/Species: Xenopus laevis (African clawed frog)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
C/Accession: S13654; S29676  
R:Gurraian, R.; Perry-O'Keefe, H.; Melton, D.A.; Weeks, D.L.  
Nature 349, 717-719, 1991  
A>Title: The Xenopus localized messenger RNA and may encode an ATP-dependent RNA helicase  
A/Reference number: S13654; MUID:9141586; PMID:1996140  
A/Accession: S13654

A:Molecule type: mRNA  
A/Residues: 1-697 <GUR>  
A/Cross-references: EMBL:X57328; NID:965059; PIDN:CAA40605.1; PID:965060  
C:Superfamily: ATP-dependent RNA helicase DBP1  
C/Keywords: ATP; nucleotide binding; P-loop  
F:260-272/Region: ATP binding #status predicted  
F:265-272/Region: nucleotide-binding motif A (P-loop)  
F:384-389/Region: nucleotide-binding motif B  
F:386-398/Region: ATP binding #status predicted  
F:388-391/Region: DEAD motif

Alignment Scores:

Pred. No.:	2,766-13	Length:	697
Score:	22.00	Matches:	22
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.04%	Indels:	0
DB:	1	Gaps:	0

US-09-714-865-15 (1-2172) x S13654 (1-697)

Qy 976 CGAGATTGATGCTTGCTCAACAGGCTGGAGAGCTGGCTTTCTCTACCA 1035  
|||||  
Db 259 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuPro 278  
|||||

Qy 1036 ATTTTG 1041  
|||||  
Db 279 I1e1eu 280

RESULT 9  
T45677  
ATP-dependent RNA helicase-like protein - Arabidopsis thaliana  
N:Alternate names: protein F14P22.160  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 17-Nov-2000  
C/Accession: T45677  
R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;  
submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z23011  
A/Accession: T45677  
A>Status: preliminary  
A:Molecule type: DNA  
A/Residues: 1-646 <DAN>  
A/Cross-references: EMBL:AL137082  
A/Experimental source: cultivar Columbia; BAC clone F14P22  
C/Genetics:  
A/Map position: 3  
A/Introns: 239/3; 267/3; 348/3; 404/3; 442/3  
A/Note: F14P22.160  
C:Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:

Pred. No.:	3,226-10	Length:	646
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.62%	Indels:	0
DB:	2	Gaps:	0

US-09-714-865-15 (1-2172) x T45677 (1-646)

Qy 970 GCAGAGCAGATTGATGCTTGCTCAACAGGCTGGAGAGCTGGCTTTCTCTACCA 1026  
|||||  
Db 182 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPhe 200  
|||||

RESULT 10  
C87818  
protein gih-1 [imported] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C/Accession: C87818  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998

A>Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog  
A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A/Accession: C87818  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-604 <STO>  
A/Cross-references: GB:chr\_I; PIDN:AAB52901.1; PID:g1947009; GSPDB:GN00019; CESP:T2IG5.3  
A/Note: T2IG5.3  
C/Genetics:  
A/Map position: 1  
A/Map position: 1  
C/Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:  
Pred. No.: 3.41e-09 Length: 604  
Score: 18.00 Matches: 18  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.49% Indels: 0  
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x C87818 (1-604)

QY 985 ATGGCTTGCTCTCAACAGGCTCTGGAGACTGCGGCTTTCTCTACCAATT 1038  
Db 223 MetacalysalaginThrglySerGlyLysThralaIaIaPheLeuLeuProIle 240

RESULT 11  
ATP-dependent RNA helicase GLH-1 - *Caenorhabditis elegans* (fragment)  
C/Species: *Caenorhabditis elegans*  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
C/Accession: T15132  
R/Minx, P.  
submitted to the EMBL Data Library, April 1997  
A/Description: The sequence of *C. elegans* cosmid T2IG5.  
A/Reference number: Z18299  
A/Accession: T15132  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-604 <MIN>  
A/Cross-references: EMBL:AF000197; NID:g1947004; PID:g1947009; PIDN:AAB52901.1; GSPDB:GN  
A/Experimental source: strain Bristol N2; clone T2IG5  
C/Genetics:  
A/Map position: 1  
A/Map position: 1

Alignment Scores:  
Pred. No.: 3.41e-09 Length: 604  
Score: 18.00 Matches: 18  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.49% Indels: 0  
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x T15132 (1-604)

QY 985 ATGGCTTGCTCTCAACAGGCTCTGGAGACTGCGGCTTTCTCTACCAATT 1038  
Db 223 MetacalysalaginThrglySerGlyLysThralaIaIaPheLeuLeuProIle 240

RESULT 12  
H84854  
probable ATP-dependent RNA helicase [imported] - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C/Accession: H84854  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: H84854  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-633 <STO>  
A/Cross-references: GB:AE002093; NID:g4559339; PIDN:AAD33001.1; GSPDB:GN00139  
A/Genetics:  
A/Map position: 2  
A/Map position: 2  
C/Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:  
Pred. No.: 3.39e-09 Length: 633  
Score: 18.00 Matches: 18  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.49% Indels: 0  
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x H84854 (1-633)

QY 973 GAGACGATTGCTGCTCAACAGGCTCTGGAGACTGCGGCTTT 1026  
Db 196 G1ArgAspLeuMetAlaCysAlaGlnThrglySerGlyLysThralaIaIaPhe 213

RESULT 13  
T45671  
ATP-dependent RNA helicase-like protein - *Arabidopsis thaliana*  
N/Alternate names: protein F14P22.100  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 17-Nov-2000  
C/Accession: T45671  
R/D'Angelo, M.; Vezzi, A.; Modesto, D.; Piazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;  
submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z23011  
A/Accession: T45671  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-603 <DN>  
A/Cross-references: EMBL:AL137082  
A/Experimental source: cultivar Columbia; BAC clone F14P22  
C/Genetics:  
A/Map position: 3  
A/Map position: 3  
A/Introns: 244/3; 272/3; 353/3; 407/3; 436/3  
A/Note: F14P22.100  
C/Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:  
Pred. No.: 3.58e-08 Length: 603  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.35% Indels: 0  
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x T45671 (1-603)

QY 976 CGAGATTGATGCTGCTCAACAGGCTCTGGAGACTGCGGCTTT 1026  
Db 189 ArgAspLeuMetAlaCysAlaGlnThrglySerGlyLysThralaIaIaPhe 205

RESULT 14  
S62003  
probable ATP-dependent RNA helicase DBP1 - yeast (*Saccharomyces cerevisiae*)  
N/Alternate names: probable RNA helicase CAL; protein LPH8c; protein YPL19c  
C/Species: *Saccharomyces cerevisiae*  
C/Date: 10-Apr-1996 #sequence\_revision 19-Apr-1996 #text\_change 19-Jan-2001  
C/Accession: S62003; S16790; A34848  
R/Schlenstedt, G.; Silver, P.A.  
submitted to the EMBL Data Library, December 1995  
A/Reference number: S61996  
A/Accession: S62003

A:Molecule type: DNA  
A:Residues: 1-617 <SCH>  
A:Cross-references: EMBL:U43503; NID:G1163087; PIDN:AAB68243.1; PID:G1163095; MIPS:YPL11  
R:Jamieson, D.J.; Beggs, J.D.  
Mol. Microbiol. 5, 805-812, 1991  
A:Title: A suppressor of yeast spp81/ded1 mutations encodes a very similar putative ATP-  
A:Reference number: S16790; MUID:91312117; PMID:1857205  
A:Accession: S16790  
A:Molecule type: DNA  
A:Residues: 1-42, 'RS', 45-47, 'K', 49-87, 'R', 89-114, 'OK', 116-617 <JAM>  
A:Cross-references: EMBL:X55993; NID:G3640; PIDN:CAA39465.1; PID:G3641  
A:Experimental source: Strain DBY939  
R:Chang, T.H.; Arenas, J.; Abelson, J.  
Proc. Natl. Acad. Sci. U.S.A. 87, 1571-1575, 1990  
A:Title: Identification of five putative yeast RNA helicase genes.  
A:Reference number: A34848; MUID:90160368; PMID:2406722  
A:Accession: A34848  
A:Molecule type: DNA  
A:Residues: 316-495, 'I', 497-500 <CHA>  
C:Genetics:  
A:Gene: SGD:DBPI  
A:Cross-references: SGD:S0006040; MIPS:YPL119C  
A:Map position: 16L  
C:Superfamily: ATP-dependent RNA helicase DBPI  
C:Keywords: ATP; nucleotide binding; P-loop  
F:198-205/Region: nucleotide-binding motif A (P-loop)  
F:314-319/Region: nucleotide-binding motif B  
F:318-321/Region: DEAD motif

Alignment Scores:  
Pred. No.: 3.57e-08 Length: 617  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.35% Indels: 0  
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x S62003 (1-617)

QY 1318 ATCAATACTTACTTTGGATGAGCTGATCGCATGTGATATGCGTTT 1368  
DB 312 IIElySTyLeuValLeuAspGIuAlaAspArgMetLeuAspMetGlyPhe 328

#### RESULT 15

T48796

probable ATP-dependent RNA helicase DED1 [Imported] - Neurospora crassa  
N:Alternate names: protein 15E6.40  
C:Species: Neurospora crassa

C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 17-Nov-2000  
C:Accession: T48796

R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24541

A:Accession: T48796

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-688 <SCH>

A:Cross-references: EMBL:ALJ53822; GSPDB:GN00112; NCSP:15E6.40

A:Experimental source: coemid contig 15E6; strain 74

C:Genetics:

A:Gene: NCSP:15E6.40

A:Map position: 2

A:Introns: 63/2

C:Superfamily: ATP-dependent RNA helicase DBPI

Alignment Scores:

Pred. No.: 3.51e-08 Length: 688  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.35% Indels: 0  
DB: 2 Gaps: 0

US-09-714-865-15 (1-2172) x T48796 (1-688)

QY 1318 ATCAATACTTACTTTGGATGAGCTGATCGCATGTGATATGCGTTT 1368  
DB 354 IIElySTyLeuValLeuAspGIuAlaAspArgMetLeuAspMetGlyPhe 370

Search completed: June 10, 2003, 17:10:06  
Job time: 73 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 17:08:46 ; Search time 42.5 Seconds

(without alignments)  
10552.370 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 724  
Sequence: 1 atgagggagatgaagatcgga.....tagatgatgagtcagtcggaat 2172

Scoring table:  
OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 392085 seqs, 103240269 residues

Word size: 1

Total number of hits satisfying chosen parameters: 738648

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Command line parameters:  
-MODE=frame+n2p\_model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US09714865/runat\_05062003\_11835\_26709/app\_query.fasta\_1.2311  
-DB=Published Applications\_AA\_QFMT=fastan\_SUFFIX=oligo.rapb\_MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=1000 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09714865 @CGN 1.1.59 @runat\_05062003\_11835\_26709  
-NCPU=6 -ICPU=3 -NO MMAR -LARGQUERY -NEG SCORES=0 -WATT -DSPELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	84	11.6	106 9	US-09-764-891-3857
2	15	2.1	238 9	US-10-195-117-3
3	15	2.1	238 9	US-10-195-117-4
4	15	2.1	343 10	US-09-925-301-1439

5	5	2.1	500 9	US-10-228-897-13	Sequence 13, Appl
6	15	2.1	575 9	US-10-108-605-55	Sequence 55, Appl
7	7	1.5	45 10	US-09-864-761-46256	Sequence 46256, A
8	11	1.5	613 10	US-09-815-242-10990	Sequence 10990, A
9	11	1.5	646 10	US-09-815-242-10311	Sequence 10311, A
10	11	1.5	646 10	US-09-815-242-13769	Sequence 13769, A
11	11	1.5	648 10	US-09-923-831-43	Sequence 43, Appl
12	10	1.4	421 10	US-09-815-242-11703	Sequence 11703, A
13	10	1.4	424 9	US-09-738-626-4359	Sequence 4359, Ap
14	10	1.4	495 9	US-09-971-536-52	Sequence 52, Appl
15	10	1.4	524 10	US-09-815-242-13491	Sequence 13491, A
16	10	1.4	536 10	US-09-815-242-10774	Sequence 10774, A
17	10	1.4	567 10	US-09-815-242-11885	Sequence 11885, A
18	10	1.4	624 10	US-09-963-7908-2	Sequence 2, Appl
19	10	1.4	732 9	US-09-738-626-4778	Sequence 4778, Ap
20	9	1.3	338 9	US-09-953-280-28	Sequence 28, Appl
21	9	1.3	338 9	US-09-953-280-29	Sequence 29, Appl
22	9	1.3	338 9	US-10-040-949A-44	Sequence 44, Appl
23	9	1.3	338 9	US-10-040-949A-45	Sequence 45, Appl
24	9	1.3	543 9	US-09-736-457-337	Sequence 337, Appl
25	9	1.3	543 9	US-09-902-941-337	Sequence 337, Appl
26	9	1.3	543 9	US-09-849-626-337	Sequence 337, Appl
27	9	1.3	543 9	US-10-017-754-337	Sequence 337, Appl
28	9	1.2	709 9	US-09-736-457-335	Sequence 335, Appl
29	9	1.2	709 9	US-09-902-941-335	Sequence 335, Appl
30	9	1.2	709 9	US-09-849-626-335	Sequence 335, Appl
31	9	1.2	709 9	US-10-017-754-335	Sequence 335, Appl
32	9	1.2	1261 12	US-10-147-268-2	Sequence 2, Appl
33	8	1.1	16 10	US-09-835-087-14	Sequence 14, Appl
34	8	1.1	39 10	US-09-864-761-42449	Sequence 42449, A
35	8	1.1	41 10	US-09-732-091-12	Sequence 12, Appl
36	8	1.1	58 10	US-10-091-504-1215	Sequence 1215, Ap
37	8	1.1	58 10	US-09-764-869-1215	Sequence 1215, Ap
38	8	1.1	64 9	US-09-764-891-4988	Sequence 488, Ap
39	8	1.1	66 9	US-09-984-271-147	Sequence 147, Appl
40	8	1.1	72 10	US-09-729-674-58	Sequence 58, Appl
41	8	1.1	95 9	US-09-984-271-234	Sequence 234, Appl
42	8	1.1	100 10	US-09-840-459-21	Sequence 21, Appl
43	8	1.1	100 10	US-09-840-459-22	Sequence 22, Appl
44	8	1.1	101 9	US-10-194-975-73	Sequence 73, Appl
45	8	1.1	101 9	US-10-194-975-74	Sequence 74, Appl
46	8	1.1	110 10	US-09-864-761-48276	Sequence 48276, A
47	8	1.1	111 9	US-09-995-529-14	Sequence 14, Appl
48	8	1.1	112 9	US-10-032-482-8	Sequence 8, Appl
49	8	1.1	112 10	US-09-772-120-6	Sequence 6, Appl
50	8	1.1	112 10	US-09-835-087-1	Sequence 1, Appl
51	8	1.1	112 10	US-09-835-087-3	Sequence 3, Appl
52	8	1.1	112 10	US-09-835-087-4	Sequence 4, Appl
53	8	1.1	112 10	US-09-835-087-5	Sequence 5, Appl
54	8	1.1	112 10	US-09-835-087-6	Sequence 6, Appl
55	8	1.1	112 10	US-09-835-087-7	Sequence 7, Appl
56	8	1.1	112 10	US-09-809-739-11	Sequence 11, Appl
57	8	1.1	112 10	US-09-809-739-14	Sequence 14, Appl
58	8	1.1	112 10	US-09-809-739-15	Sequence 15, Appl
59	8	1.1	112 10	US-09-809-739-16	Sequence 16, Appl
60	8	1.1	112 10	US-09-809-739-17	Sequence 17, Appl
61	8	1.1	112 10	US-09-809-739-18	Sequence 18, Appl
62	8	1.1	112 10	US-09-840-459-9	Sequence 9, Appl
63	8	1.1	112 10	US-09-840-459-12	Sequence 12, Appl
64	8	1.1	112 10	US-09-840-459-13	Sequence 13, Appl
65	8	1.1	112 10	US-09-840-459-14	Sequence 14, Appl
66	8	1.1	112 10	US-09-840-459-15	Sequence 15, Appl
67	8	1.1	112 10	US-09-840-459-15	Sequence 15, Appl
68	8	1.1	112 10	US-09-840-459-107	Sequence 107, Appl
69	8	1.1	113 9	US-09-995-529-18	Sequence 18, Appl
70	8	1.1	113 9	US-09-840-459-68	Sequence 68, Appl
71	8	1.1	114 10	US-09-840-459-106	Sequence 106, Appl
72	8	1.1	133 10	US-09-925-301-1176	Sequence 1176, Ap
73	8	1.1	142 9	US-09-840-459-1102	Sequence 102, Appl
74	8	1.1	160 9	US-09-991-496-135	Sequence 135, Appl
75	8	1.1	211 10	US-09-825-297-823	Sequence 823, Appl
76	8	1.1	220 9	US-09-925-299-843	Sequence 843, Appl
77	8	1.1	220 10	US-09-925-299-843	Sequence 843, Appl

C 78	8	1.1	233	10	US-09-938-803-15	Sequence 15, App1	151	7	1.0	76	9	US-09-895-814-888	Sequence 888, App
C 79	8	1.1	236	10	US-09-925-300-1410	Sequence 1410, App	152	7	1.0	76	10	US-09-759-143-575	Sequence 575, App
C 80	8	1.1	239	9	US-10-195-117-6	Sequence 6, App1	153	7	1.0	76	10	US-09-759-143-575	Sequence 888, App
C 81	8	1.1	271	10	US-09-911-346-2	Sequence 2, App1	154	7	1.0	76	10	US-09-780-669-575	Sequence 575, App
C 82	8	1.1	320	9	US-10-278-173-66	Sequence 68, App1	155	7	1.0	76	10	US-09-822-827-575	Sequence 888, App
C 83	8	1.1	328	10	US-09-815-242-1348	Sequence 1348, A	156	7	1.0	76	10	US-09-822-827-575	Sequence 575, App
C 84	8	1.1	328	10	US-09-815-242-13556	Sequence 13556, A	157	7	1.0	76	10	US-09-822-827-575	Sequence 888, App
C 85	8	1.1	353	9	US-09-953-280-42	Sequence 42, App1	158	7	1.0	77	10	US-09-864-761-37461	Sequence 37461, A
C 86	8	1.1	353	9	US-10-040-949A-58	Sequence 58, App1	159	7	1.0	80	10	US-09-864-761-37461	Sequence 46876, A
C 87	8	1.1	387	10	US-09-925-900-4477	Sequence 1477, App	160	7	1.0	86	10	US-09-864-761-34118	Sequence 34118, A
C 88	8	1.1	440	9	US-09-738-626-6376	Sequence 6376, App	161	7	1.0	88	10	US-09-939-980-512	Sequence 513, App
C 89	8	1.1	459	10	US-09-925-501-1027	Sequence 1027, App	162	7	1.0	89	9	US-09-764-891-1515	Sequence 1517, App
C 90	8	1.1	471	9	US-10-082-830-265	Sequence 265, App	163	7	1.0	95	10	US-09-864-761-48467	Sequence 48467, A
C 91	8	1.1	487	10	US-09-815-242-12397	Sequence 12397, A	164	7	1.0	97	10	US-09-734-0178-40	Sequence 40, App1
C 92	8	1.1	487	10	US-09-815-242-12991	Sequence 12991, A	165	7	1.0	98	10	US-09-864-761-34193	Sequence 34193, A
C 93	8	1.1	492	10	US-09-815-242-11308	Sequence 11308, A	166	7	1.0	99	10	US-09-815-242-5533	Sequence 5533, App
C 94	8	1.1	492	10	US-09-732-091-2	Sequence 2, App1	167	7	1.0	101	9	US-10-097-665-366	Sequence 366, App
C 95	8	1.1	504	10	US-09-732-091-42	Sequence 42, App1	168	7	1.0	101	9	US-09-986-460-331	Sequence 231, App
C 96	8	1.1	510	10	US-09-815-242-5256	Sequence 5256, App	169	7	1.0	105	10	US-09-815-242-12217	Sequence 12217, A
C 97	8	1.1	526	9	US-10-084-205-74	Sequence 74, App1	170	7	1.0	105	10	US-09-815-242-12760	Sequence 12760, A
C 98	8	1.1	535	10	US-09-925-637-74	Sequence 74, App1	171	7	1.0	111	10	US-09-893-737-122	Sequence 122, App
C 99	8	1.1	548	9	US-09-968-851-38	Sequence 38, App1	172	7	1.0	111	10	US-09-746-491-16	Sequence 16, App1
C 100	8	1.1	548	9	US-10-023-282-264	Sequence 264, App	173	7	1.0	112	9	US-09-898-751A-12	Sequence 12, App1
C 101	8	1.1	633	9	US-09-834-998A-3	Sequence 3, App1	174	7	1.0	112	9	US-10-146-496-6	Sequence 6, App1
C 102	8	1.1	725	9	US-10-108-605-321	Sequence 321, App	175	7	1.0	112	10	US-09-931-382A-11	Sequence 11, App1
C 103	8	1.1	752	10	US-09-825-809-2	Sequence 2, App1	176	7	1.0	114	9	US-10-114-893-137	Sequence 137, App
C 104	8	1.1	752	10	US-09-825-809-4	Sequence 4, App1	177	7	1.0	117	9	US-10-193-616-6	Sequence 6, App1
C 105	8	1.1	764	10	US-09-902-627-1	Sequence 1, App1	178	7	1.0	119	9	US-10-050-704-227	Sequence 227, App
C 106	8	1.1	805	9	US-10-108-605-113	Sequence 113, App	179	7	1.0	122	10	US-09-893-737-286	Sequence 286, App
C 107	8	1.1	1345	9	US-10-108-605-249	Sequence 249, App	180	7	1.0	123	9	US-09-796-692-1360	Sequence 1360, App
C 108	8	1.1	1661	9	US-09-842-758-42	Sequence 42, App1	181	7	1.0	123	9	US-09-796-692-1360	Sequence 229, App
C 109	8	1.1	1744	9	US-10-108-605-25	Sequence 25, App1	182	7	1.0	123	9	US-10-040-862-1360	Sequence 1360, App
C 110	7	1.0	10	9	US-09-572-404B-1213	Sequence 1213, App	183	7	1.0	123	9	US-10-040-862-1322	Sequence 229, App
C 111	7	1.0	12	9	US-10-057-789-305	Sequence 305, App	184	7	1.0	130	9	US-09-948-783-142	Sequence 142, App
C 112	7	1.0	12	9	US-10-212-628-305	Sequence 305, App	185	7	1.0	131	9	US-09-892-877-140	Sequence 140, App
C 113	7	1.0	12	9	US-09-954-385-134	Sequence 134, App	186	7	1.0	135	10	US-10-050-704-107	Sequence 107, App
C 114	7	1.0	14	9	US-09-984-245-147	Sequence 147, App	187	7	1.0	136	9	US-09-764-891-4420	Sequence 4420, App
C 115	7	1.0	14	9	US-09-966-262-147	Sequence 147, App	188	7	1.0	137	9	US-09-764-891-4420	Sequence 147, App
C 116	7	1.0	14	9	US-09-983-966-147	Sequence 147, App	189	7	1.0	138	9	US-10-042-141-151	Sequence 151, App
C 117	7	1.0	14	9	US-10-143-090-147	Sequence 147, App	190	7	1.0	138	10	US-09-726-643-151	Sequence 151, App
C 118	7	1.0	18	9	US-09-753-130B-294	Sequence 294, App	191	7	1.0	141	10	US-09-893-737-254	Sequence 254, App
C 119	7	1.0	18	9	US-10-189-123-24	Sequence 24, App1	192	7	1.0	142	9	US-09-966-546-22	Sequence 22, App1
C 120	7	1.0	18	10	US-09-864-761-40388	Sequence 40388, A	193	7	1.0	142	9	US-09-966-545-22	Sequence 22, App1
C 121	7	1.0	19	10	US-09-879-957-170	Sequence 170, App	194	7	1.0	142	9	US-09-965-212-22	Sequence 22, App1
C 122	7	1.0	22	9	US-10-097-065-369	Sequence 369, App	195	7	1.0	142	10	US-09-800-729-139	Sequence 139, App
C 123	7	1.0	22	10	US-09-071-838-302	Sequence 302, App	196	7	1.0	143	10	US-09-800-729-139	Sequence 198, App
C 124	7	1.0	22	10	US-09-864-761-33674	Sequence 33674, A	197	7	1.0	143	9	US-09-782-974C-62	Sequence 62, App1
C 125	7	1.0	33	10	US-10-042-141-155	Sequence 155, App	198	7	1.0	150	10	US-09-893-737-230	Sequence 230, App
C 126	7	1.0	33	10	US-09-864-761-48994	Sequence 48994, A	199	7	1.0	152	9	US-09-764-891-1604	Sequence 4604, App
C 127	7	1.0	33	10	US-09-726-643-155	Sequence 155, App	200	7	1.0	155	9	US-09-738-626-4054	Sequence 4054, App
C 128	7	1.0	34	10	US-09-864-761-34605	Sequence 34605, A	201	7	1.0	157	9	US-10-080-960-27	Sequence 27, App1
C 129	7	1.0	37	10	US-09-864-761-39859	Sequence 39859, A	202	7	1.0	161	9	US-10-099-766-2	Sequence 2, App1
C 130	7	1.0	37	10	US-09-864-761-42324	Sequence 42324, A	203	7	1.0	161	10	US-09-682-706-5	Sequence 5, App1
C 131	7	1.0	41	10	US-09-864-761-37109	Sequence 37109, A	204	7	1.0	168	9	US-10-267-718-19	Sequence 19, App1
C 132	7	1.0	43	9	US-09-981-876-205	Sequence 205, App	205	7	1.0	168	10	US-09-354-453-19	Sequence 19, App1
C 133	7	1.0	43	9	US-09-148-545-205	Sequence 205, App	206	7	1.0	175	9	US-10-050-704-428	Sequence 228, App
C 134	7	1.0	45	10	US-10-072-349-102	Sequence 102, App	207	7	1.0	176	9	US-10-193-616-8	Sequence 8, App1
C 135	7	1.0	45	10	US-09-764-855-102	Sequence 102, App	208	7	1.0	176	10	US-09-855-266A-1	Sequence 4, App1
C 136	7	1.0	48	9	US-09-984-271-165	Sequence 165, App	209	7	1.0	177	10	US-09-755-325-4	Sequence 5, App1
C 137	7	1.0	53	9	US-10-150-111-81	Sequence 81, App1	210	7	1.0	180	9	US-09-989-442-107	Sequence 107, App
C 138	7	1.0	56	9	US-10-114-893-44	Sequence 44, App1	211	7	1.0	180	9	US-10-073-865-86	Sequence 86, App1
C 139	7	1.0	56	9	US-10-114-893-129	Sequence 129, App	212	7	1.0	180	9	US-10-103-313-179	Sequence 319, App
C 140	7	1.0	60	10	US-09-864-761-40431	Sequence 40431, A	213	7	1.0	180	9	US-10-073-885-99	Sequence 99, App1
C 141	7	1.0	61	9	US-09-983-802-621	Sequence 621, App	214	7	1.0	180	10	US-09-764-853-567	Sequence 567, App
C 142	7	1.0	64	10	US-09-922-261-267	Sequence 267, App	215	7	1.0	183	9	US-09-989-442-111	Sequence 111, App
C 143	7	1.0	65	10	US-09-867-550-54	Sequence 54, App1	216	7	1.0	183	9	US-10-073-865-95	Sequence 95, App1
C 144	7	1.0	72	9	US-09-746-783-118	Sequence 118, App	217	7	1.0	183	9	US-10-103-313-189	Sequence 189, App
C 145	7	1.0	75	10	US-09-932-261-265	Sequence 265, App	218	7	1.0	183	9	US-10-073-885-102	Sequence 102, App
C 146	7	1.0	76	9	US-10-012-896-575	Sequence 575, App	219	7	1.0	183	10	US-09-764-853-592	Sequence 592, App
C 147	7	1.0	76	9	US-10-012-896-888	Sequence 888, App	220	7	1.0	184	10	US-09-925-301-847	Sequence 847, App
C 148	7	1.0	76	9	US-09-895-793-575	Sequence 575, App	221	7	1.0	185	10	US-09-529-063-34	Sequence 34, App1
C 149	7	1.0	76	9	US-09-895-793-888	Sequence 888, App	222	7	1.0	188	9	US-09-866-050A-686	Sequence 686, App
C 150	7	1.0	76	9	US-09-895-814-575	Sequence 575, App	223	7	1.0	198	9	US-09-892-877-112	Sequence 212, App



224	7	1.0	198	9	US-09-948-783-213	Sequence 213, App	297	7	1.0	311	9	US-10-189-123-11	Sequence 11, Appl
225	7	1.0	203	9	US-10-050-704-244	Sequence 244, Appl	298	7	1.0	333	10	US-10-318-142-6	Sequence 6, Appl1
226	7	1.0	210	9	US-09-973-025-14	Sequence 14, Appl	299	7	1.0	373	10	US-09-760-352-A-2	Sequence 2, Appl
227	7	1.0	210	9	US-09-899-303-14	Sequence 14, Appl	300	7	1.0	373	10	US-09-820-899-55	Sequence 55, Appl
228	7	1.0	210	9	US-09-995-808-14	Sequence 14, Appl	301	7	1.0	375	10	US-09-764-864-1316	Sequence 1316, Ap
229	7	1.0	215	9	US-10-255-532-4	Sequence 4, Appl1	302	7	1.0	378	9	US-10-073-885-77	Sequence 77, Appl
230	7	1.0	223	9	US-09-875-921-9	Sequence 9, Appl1	303	7	1.0	378	10	US-09-764-864-863	Sequence 863, App
231	7	1.0	223	9	US-10-162-012-9	Sequence 9, Appl1	304	7	1.0	382	12	US-10-078-923-34	Sequence 34, Appl
232	7	1.0	228	10	US-09-780-317-29	Sequence 29, Appl	305	7	1.0	383	9	US-09-738-626-6821	Sequence 6821, Ap
233	7	1.0	229	10	US-09-804-357-12	Sequence 12, Appl	306	7	1.0	386	9	US-10-270-333-18	Sequence 18, Appl
234	7	1.0	229	10	US-09-804-006-12	Sequence 12, Appl	307	7	1.0	387	9	US-10-114-893-133	Sequence 133, App
235	7	1.0	242	10	US-09-864-761-36180	Sequence 36180, A	308	7	1.0	387	9	US-10-016-249-2	Sequence 2, Appl1
236	7	1.0	242	10	US-09-867-550-1878	Sequence 1878, Ap	309	7	1.0	383	10	US-09-823-240-11	Sequence 11, Appl
237	7	1.0	245	9	US-10-128-870-8	Sequence 8, Appl1	310	7	1.0	386	9	US-10-193-616-14	Sequence 10, Appl
238	7	1.0	245	9	US-10-131-685-8	Sequence 8, Appl1	311	7	1.0	399	10	US-09-893-737-10	Sequence 10, Appl
239	7	1.0	245	10	US-09-942-858-9	Sequence 9, Appl1	312	7	1.0	402	10	US-09-764-864-1496	Sequence 1496, Ap
240	7	1.0	247	10	US-09-893-348-14	Sequence 14, Appl	313	7	1.0	403	9	US-09-991-496-10	Sequence 10, Appl
241	7	1.0	248	9	US-10-233-873A-3	Sequence 3, Appl1	314	7	1.0	403	10	US-09-874-925-10	Sequence 10, Appl
242	7	1.0	257	12	US-08-450-842-10	Sequence 10, Appl	315	7	1.0	404	9	US-09-712-363-257	Sequence 257, App
243	7	1.0	259	12	US-10-078-929-197	Sequence 197, App	316	7	1.0	404	10	US-09-764-864-1074	Sequence 1074, Ap
244	7	1.0	259	10	US-09-911-826A-6	Sequence 6, Appl1	317	7	1.0	407	9	US-09-770-509-16	Sequence 16, Appl
245	7	1.0	268	10	US-09-864-866-43	Sequence 43, Appl	318	7	1.0	408	9	US-09-988-462-21	Sequence 21, Appl
246	7	1.0	272	9	US-10-102-806-663	Sequence 663, App	319	7	1.0	410	10	US-09-925-300-1074	Sequence 1074, Ap
247	7	1.0	275	9	US-10-112-645-4	Sequence 4, Appl1	320	7	1.0	411	1	US-08-954-771-10	Sequence 10, Appl
248	7	1.0	277	9	US-10-101-464A-631	Sequence 631, App	321	7	1.0	411	8	US-08-900-220C-12	Sequence 12, Appl
249	7	1.0	284	9	US-10-073-885-71	Sequence 71, Appl	322	7	1.0	411	9	US-09-883-848A-12	Sequence 12, Appl
250	7	1.0	284	10	US-09-771-161A-106	Sequence 106, App	323	7	1.0	411	9	US-09-187-387-12	Sequence 12, Appl
251	7	1.0	290	9	US-10-117-323-4	Sequence 4, Appl1	324	7	1.0	411	10	US-09-755-325-6	Sequence 6, Appl1
252	7	1.0	292	9	US-10-278-173-18	Sequence 18, Appl	325	7	1.0	411	10	US-09-755-325-8	Sequence 8, Appl1
253	7	1.0	292	10	US-09-864-761-37944	Sequence 37944, A	326	7	1.0	411	10	US-09-151-999-12	Sequence 12, Appl
254	7	1.0	293	10	US-09-815-242-5614	Sequence 5614, Ap	327	7	1.0	415	10	US-09-755-325-2	Sequence 2, Appl1
255	7	1.0	298	10	US-09-789-919-68	Sequence 68, App	328	7	1.0	417	9	US-10-075-846-2	Sequence 2, Appl1
256	7	1.0	299	9	US-10-125-001-22	Sequence 22, Appl	329	7	1.0	431	9	US-10-075-846-4	Sequence 4, Appl1
257	7	1.0	300	9	US-10-128-870-4	Sequence 4, Appl1	330	7	1.0	434	10	US-09-801-366-146	Sequence 146, App
258	7	1.0	300	9	US-10-128-870-6	Sequence 6, Appl1	331	7	1.0	437	9	US-09-941-947A-2	Sequence 2, Appl1
259	7	1.0	300	9	US-10-131-685-4	Sequence 4, Appl1	332	7	1.0	437	10	US-09-815-242-5523	Sequence 523, Ap
260	7	1.0	300	9	US-10-131-685-6	Sequence 6, Appl1	333	7	1.0	437	10	US-09-934-901-16	Sequence 16, Appl
261	7	1.0	300	10	US-09-861-451A-42	Sequence 42, Appl	334	7	1.0	437	10	US-09-934-901-16	Sequence 6, Appl1
262	7	1.0	301	9	US-10-004-633-24	Sequence 24, Appl	335	7	1.0	440	9	US-10-063-547-134	Sequence 134, App
263	7	1.0	301	9	US-10-183-116-71	Sequence 71, Appl	336	7	1.0	440	9	US-10-174-590-386	Sequence 386, App
264	7	1.0	303	9	US-09-828-523A-20	Sequence 20, Appl	337	7	1.0	440	9	US-10-176-758-386	Sequence 386, App
265	7	1.0	303	9	US-09-533-029-6	Sequence 6, Appl1	338	7	1.0	440	9	US-10-063-616-134	Sequence 134, App
266	7	1.0	303	10	US-09-815-242-12527	Sequence 12527, A	339	7	1.0	440	9	US-10-175-737-386	Sequence 386, App
267	7	1.0	308	9	US-10-097-065-150	Sequence 150, App	340	7	1.0	440	9	US-10-063-502-134	Sequence 134, App
268	7	1.0	309	9	US-09-991-496-134	Sequence 134, App	341	7	1.0	440	9	US-10-173-706-386	Sequence 386, App
269	7	1.0	311	9	US-09-828-523A-86	Sequence 86, Appl	342	7	1.0	440	9	US-10-173-706-386	Sequence 386, App
270	7	1.0	313	9	US-09-771-503-9	Sequence 9, Appl1	343	7	1.0	440	9	US-10-175-732-386	Sequence 386, App
271	7	1.0	326	10	US-09-804-357-8	Sequence 8, Appl1	344	7	1.0	440	9	US-10-176-482-386	Sequence 386, App
272	7	1.0	326	10	US-09-804-006-8	Sequence 8, Appl1	345	7	1.0	440	9	US-10-176-757-386	Sequence 386, App
273	7	1.0	332	9	US-10-236-250-2	Sequence 2, Appl1	346	7	1.0	440	9	US-10-176-813-386	Sequence 386, App
274	7	1.0	332	9	US-09-804-291-63	Sequence 63, Appl	347	7	1.0	440	9	US-10-180-552-386	Sequence 386, App
275	7	1.0	332	10	US-09-767-041-21	Sequence 21, Appl	348	7	1.0	440	9	US-10-180-557-386	Sequence 386, App
276	7	1.0	332	10	US-09-886-055-63	Sequence 63, Appl	349	7	1.0	440	9	US-10-173-700-386	Sequence 386, App
277	7	1.0	332	10	US-09-764-864-1573	Sequence 1573, Ap	350	7	1.0	440	9	US-10-174-572-386	Sequence 386, App
278	7	1.0	332	10	US-09-771-730-10	Sequence 10, Appl	351	7	1.0	440	9	US-10-174-579-386	Sequence 386, App
279	7	1.0	332	10	US-09-771-730-28	Sequence 28, Appl	352	7	1.0	440	9	US-10-174-582-386	Sequence 386, App
280	7	1.0	340	9	US-09-738-626-5037	Sequence 5037, Ap	353	7	1.0	440	9	US-10-174-588-386	Sequence 386, App
281	7	1.0	344	9	US-09-866-582-6	Sequence 6, Appl1	354	7	1.0	440	9	US-10-175-739-386	Sequence 386, App
282	7	1.0	348	10	US-09-815-242-5150	Sequence 5150, Ap	355	7	1.0	440	9	US-10-175-740-386	Sequence 386, App
283	7	1.0	353	9	US-09-759-130B-283	Sequence 283, App	356	7	1.0	440	9	US-10-175-743-386	Sequence 386, App
284	7	1.0	353	9	US-10-189-123-13	Sequence 13, Appl	357	7	1.0	440	9	US-10-176-488-386	Sequence 386, App
285	7	1.0	359	9	US-10-093-246-1	Sequence 1, Appl1	358	7	1.0	440	9	US-10-176-492-386	Sequence 386, App
286	7	1.0	359	12	US-10-093-045-1	Sequence 1, Appl1	359	7	1.0	440	9	US-10-176-492-386	Sequence 386, App
287	7	1.0	360	10	US-09-925-301-1116	Sequence 1116, Ap	360	7	1.0	440	9	US-10-176-747-386	Sequence 386, App
288	7	1.0	361	10	US-09-796-858-28	Sequence 28, Appl	361	7	1.0	440	9	US-10-176-750-386	Sequence 386, App
289	7	1.0	362	10	US-09-933-246-2	Sequence 2, Appl1	362	7	1.0	440	9	US-10-176-985-386	Sequence 386, App
290	7	1.0	362	12	US-10-093-045-2	Sequence 2, Appl1	363	7	1.0	440	9	US-10-176-987-386	Sequence 386, App
291	7	1.0	366	10	US-09-934-899-6	Sequence 6, Appl1	364	7	1.0	440	9	US-10-176-991-386	Sequence 386, App
292	7	1.0	366	10	US-09-934-868-26	Sequence 26, Appl	365	7	1.0	440	9	US-10-176-992-386	Sequence 386, App
293	7	1.0	370	10	US-10-097-065-381	Sequence 381, App	366	7	1.0	440	9	US-10-184-658-386	Sequence 386, App
294	7	1.0	370	10	US-09-893-737-202	Sequence 202, App	367	7	1.0	440	9	US-10-173-695-386	Sequence 386, App
295	7	1.0	371	9	US-09-923-844B-2	Sequence 2, Appl1	368	7	1.0	440	9	US-10-173-697-386	Sequence 386, App
296	7	1.0	371	9	US-09-759-130B-281	Sequence 281, App	369	7	1.0	440	9	US-10-173-705-386	Sequence 386, App

C 370	7	1.0	440	9	US-10-174-576-386	Sequence 386, App	C 443	7	1.0	440	9	US-10-187-757-386	Sequence 386, App
C 371	7	1.0	440	9	US-10-174-586-386	Sequence 386, App	C 444	7	1.0	440	9	US-10-187-767-386	Sequence 386, App
C 372	7	1.0	440	9	US-10-174-586-386	Sequence 386, App	C 445	7	1.0	440	9	US-10-187-767-386	Sequence 386, App
C 373	7	1.0	440	9	US-10-175-747-386	Sequence 386, App	C 446	7	1.0	440	9	US-10-188-769-386	Sequence 386, App
C 374	7	1.0	440	9	US-10-175-747-386	Sequence 386, App	C 447	7	1.0	440	9	US-10-188-769-386	Sequence 386, App
C 375	7	1.0	440	9	US-10-176-485-386	Sequence 386, App	C 448	7	1.0	440	9	US-10-188-770-386	Sequence 386, App
C 376	7	1.0	440	9	US-10-176-485-386	Sequence 386, App	C 449	7	1.0	440	9	US-10-188-770-386	Sequence 386, App
C 377	7	1.0	440	9	US-10-176-493-386	Sequence 386, App	C 450	7	1.0	440	9	US-10-188-781-386	Sequence 386, App
C 378	7	1.0	440	9	US-10-176-756-386	Sequence 386, App	C 451	7	1.0	440	9	US-10-194-361-386	Sequence 386, App
C 379	7	1.0	440	9	US-10-176-911-386	Sequence 386, App	C 452	7	1.0	440	9	US-10-194-423-386	Sequence 386, App
C 380	7	1.0	440	9	US-10-176-919-386	Sequence 386, App	C 453	7	1.0	440	9	US-10-195-897-386	Sequence 386, App
C 381	7	1.0	440	9	US-10-176-925-386	Sequence 386, App	C 454	7	1.0	440	9	US-10-195-901-386	Sequence 386, App
C 382	7	1.0	440	9	US-10-176-978-386	Sequence 386, App	C 455	7	1.0	440	9	US-10-196-756-386	Sequence 386, App
C 383	7	1.0	440	9	US-10-179-510-386	Sequence 386, App	C 456	7	1.0	440	9	US-10-179-708-386	Sequence 386, App
C 384	7	1.0	440	9	US-10-180-543-386	Sequence 386, App	C 457	7	1.0	440	9	US-10-176-479-386	Sequence 386, App
C 385	7	1.0	440	9	US-10-180-544-386	Sequence 386, App	C 458	7	1.0	440	9	US-10-176-748-386	Sequence 386, App
C 386	7	1.0	440	9	US-10-180-546-386	Sequence 386, App	C 459	7	1.0	440	9	US-10-176-916-386	Sequence 386, App
C 387	7	1.0	440	9	US-10-180-547-386	Sequence 386, App	C 460	7	1.0	440	9	US-10-179-507-386	Sequence 386, App
C 388	7	1.0	440	9	US-10-180-548-386	Sequence 386, App	C 461	7	1.0	440	9	US-10-179-516-386	Sequence 386, App
C 389	7	1.0	440	9	US-10-180-555-386	Sequence 386, App	C 462	7	1.0	440	9	US-10-179-519-386	Sequence 386, App
C 390	7	1.0	440	9	US-10-180-559-386	Sequence 386, App	C 463	7	1.0	440	9	US-10-179-525-386	Sequence 386, App
C 391	7	1.0	440	9	US-10-181-000-386	Sequence 386, App	C 464	7	1.0	440	9	US-10-180-540-386	Sequence 386, App
C 392	7	1.0	440	9	US-10-183-010-386	Sequence 386, App	C 465	7	1.0	440	9	US-10-180-545-386	Sequence 386, App
C 393	7	1.0	440	9	US-10-183-012-386	Sequence 386, App	C 466	7	1.0	440	9	US-10-183-006-386	Sequence 386, App
C 394	7	1.0	440	9	US-10-184-614-386	Sequence 386, App	C 467	7	1.0	440	9	US-10-183-008-386	Sequence 386, App
C 395	7	1.0	440	9	US-10-184-623-386	Sequence 386, App	C 468	7	1.0	440	9	US-10-183-017-386	Sequence 386, App
C 396	7	1.0	440	9	US-10-184-635-386	Sequence 386, App	C 469	7	1.0	440	9	US-10-183-019-386	Sequence 386, App
C 397	7	1.0	440	9	US-10-184-637-386	Sequence 386, App	C 470	7	1.0	440	9	US-10-184-618-386	Sequence 386, App
C 398	7	1.0	440	9	US-10-184-646-386	Sequence 386, App	C 471	7	1.0	440	9	US-10-184-625-386	Sequence 386, App
C 399	7	1.0	440	9	US-10-184-647-386	Sequence 386, App	C 472	7	1.0	440	9	US-10-184-626-386	Sequence 386, App
C 400	7	1.0	440	9	US-10-184-652-386	Sequence 386, App	C 473	7	1.0	440	9	US-10-184-627-386	Sequence 386, App
C 401	7	1.0	440	9	US-10-187-594-386	Sequence 386, App	C 474	7	1.0	440	9	US-10-184-645-386	Sequence 386, App
C 402	7	1.0	440	9	US-10-187-596-386	Sequence 386, App	C 475	7	1.0	440	9	US-10-184-654-386	Sequence 386, App
C 403	7	1.0	440	9	US-10-187-747-386	Sequence 386, App	C 476	7	1.0	440	9	US-10-184-655-386	Sequence 386, App
C 404	7	1.0	440	9	US-10-187-885-386	Sequence 386, App	C 477	7	1.0	440	9	US-10-188-774-386	Sequence 386, App
C 405	7	1.0	440	9	US-10-187-886-386	Sequence 386, App	C 478	7	1.0	440	9	US-10-188-775-386	Sequence 386, App
C 406	7	1.0	440	9	US-10-199-464-386	Sequence 386, App	C 479	7	1.0	440	9	US-10-194-462-386	Sequence 386, App
C 407	7	1.0	440	9	US-10-176-751-386	Sequence 386, App	C 480	7	1.0	440	9	US-10-199-902-386	Sequence 386, App
C 408	7	1.0	440	9	US-10-176-760-386	Sequence 386, App	C 481	7	1.0	440	9	US-10-196-743-386	Sequence 386, App
C 409	7	1.0	440	9	US-10-176-990-386	Sequence 386, App	C 482	7	1.0	440	9	US-10-196-745-386	Sequence 386, App
C 410	7	1.0	440	9	US-10-180-541-386	Sequence 386, App	C 483	7	1.0	440	9	US-10-196-760-386	Sequence 386, App
C 411	7	1.0	440	9	US-10-180-542-386	Sequence 386, App	C 484	7	1.0	440	9	US-10-196-762-386	Sequence 386, App
C 412	7	1.0	440	9	US-10-180-548-386	Sequence 386, App	C 485	7	1.0	440	9	US-10-197-695-386	Sequence 386, App
C 413	7	1.0	440	9	US-10-180-551-386	Sequence 386, App	C 486	7	1.0	440	9	US-10-176-484-386	Sequence 386, App
C 414	7	1.0	440	9	US-10-180-998-386	Sequence 386, App	C 487	7	1.0	440	9	US-10-176-753-386	Sequence 386, App
C 415	7	1.0	440	9	US-10-180-999-386	Sequence 386, App	C 488	7	1.0	440	9	US-10-176-917-386	Sequence 386, App
C 416	7	1.0	440	9	US-10-183-013-386	Sequence 386, App	C 489	7	1.0	440	9	US-10-176-982-386	Sequence 386, App
C 417	7	1.0	440	9	US-10-184-612-386	Sequence 386, App	C 490	7	1.0	440	9	US-10-179-506-386	Sequence 386, App
C 418	7	1.0	440	9	US-10-184-616-386	Sequence 386, App	C 491	7	1.0	440	9	US-10-179-513-386	Sequence 386, App
C 419	7	1.0	440	9	US-10-184-617-386	Sequence 386, App	C 492	7	1.0	440	9	US-10-179-514-386	Sequence 386, App
C 420	7	1.0	440	9	US-10-184-622-386	Sequence 386, App	C 493	7	1.0	440	9	US-10-179-522-386	Sequence 386, App
C 421	7	1.0	440	9	US-10-184-628-386	Sequence 386, App	C 494	7	1.0	440	9	US-10-180-556-386	Sequence 386, App
C 422	7	1.0	440	9	US-10-184-629-386	Sequence 386, App	C 495	7	1.0	440	9	US-10-180-560-386	Sequence 386, App
C 423	7	1.0	440	9	US-10-184-630-386	Sequence 386, App	C 496	7	1.0	440	9	US-10-183-015-386	Sequence 386, App
C 424	7	1.0	440	9	US-10-184-631-386	Sequence 386, App	C 497	7	1.0	440	9	US-10-184-615-386	Sequence 386, App
C 425	7	1.0	440	9	US-10-184-632-386	Sequence 386, App	C 498	7	1.0	440	9	US-10-184-620-386	Sequence 386, App
C 426	7	1.0	440	9	US-10-184-636-386	Sequence 386, App	C 499	7	1.0	440	9	US-10-184-643-386	Sequence 386, App
C 427	7	1.0	440	9	US-10-184-640-386	Sequence 386, App	C 500	7	1.0	440	9	US-10-184-656-386	Sequence 386, App
C 428	7	1.0	440	9	US-10-184-650-386	Sequence 386, App	C 501	7	1.0	440	9	US-10-192-010-386	Sequence 386, App
C 429	7	1.0	440	9	US-10-184-651-386	Sequence 386, App	C 502	7	1.0	440	9	US-10-195-898-386	Sequence 386, App
C 430	7	1.0	440	9	US-10-187-588-386	Sequence 386, App	C 503	7	1.0	440	9	US-10-209-908-386	Sequence 386, App
C 431	7	1.0	440	9	US-10-187-597-386	Sequence 386, App	C 504	7	1.0	440	9	US-10-063-518-134	Sequence 134, App
C 432	7	1.0	440	9	US-10-187-598-386	Sequence 386, App	C 505	7	1.0	440	9	US-10-063-598-134	Sequence 386, App
C 433	7	1.0	440	9	US-10-187-600-386	Sequence 386, App	C 506	7	1.0	440	9	US-10-184-619-386	Sequence 386, App
C 434	7	1.0	440	9	US-10-187-601-386	Sequence 386, App	C 507	7	1.0	440	9	US-10-186-855-386	Sequence 386, App
C 435	7	1.0	440	9	US-10-187-602-386	Sequence 386, App	C 508	7	1.0	440	9	US-10-187-599-386	Sequence 386, App
C 436	7	1.0	440	9	US-10-187-741-386	Sequence 386, App	C 509	7	1.0	440	9	US-10-187-750-386	Sequence 386, App
C 437	7	1.0	440	9	US-10-187-743-386	Sequence 386, App	C 510	7	1.0	440	9	US-10-188-760-386	Sequence 386, App
C 438	7	1.0	440	9	US-10-187-743-386	Sequence 386, App	C 511	7	1.0	440	9	US-10-192-015-386	Sequence 386, App
C 439	7	1.0	440	9	US-10-187-746-386	Sequence 386, App	C 512	7	1.0	440	9	US-10-194-394-386	Sequence 386, App
C 440	7	1.0	440	9	US-10-187-747-386	Sequence 386, App	C 513	7	1.0	440	9	US-10-194-425-386	Sequence 386, App
C 441	7	1.0	440	9	US-10-187-751-386	Sequence 386, App	C 514	7	1.0	440	9	US-10-194-465-386	Sequence 386, App
C 442	7	1.0	440	9	US-10-187-753-386	Sequence 386, App	C 515	7	1.0	440	9	US-10-195-885-386	Sequence 386, App

C 516	7	1.0	440	9	US-10-195-890-386	Sequence 386, App	C 589	7	1.0	440	9	US-10-200-617-386	Sequence 386, App
C 517	7	1.0	440	9	US-10-195-893-386	Sequence 386, App	C 590	7	1.0	440	9	US-10-205-893-386	Sequence 386, App
C 518	7	1.0	440	9	US-10-196-748-386	Sequence 386, App	C 591	7	1.0	440	9	US-10-205-897-386	Sequence 386, App
C 519	7	1.0	440	9	US-10-196-750-386	Sequence 386, App	C 592	7	1.0	440	9	US-10-063-567-134	Sequence 134, App
C 520	7	1.0	440	9	US-10-197-699-386	Sequence 386, App	C 593	7	1.0	440	9	US-10-174-571-386	Sequence 386, App
C 521	7	1.0	440	9	US-10-197-700-386	Sequence 386, App	C 594	7	1.0	440	9	US-10-176-746-386	Sequence 386, App
C 522	7	1.0	440	9	US-10-197-705-386	Sequence 386, App	C 595	7	1.0	440	9	US-10-176-923-386	Sequence 386, App
C 523	7	1.0	440	9	US-10-197-708-386	Sequence 386, App	C 596	7	1.0	440	9	US-10-183-011-386	Sequence 386, App
C 524	7	1.0	440	9	US-10-198-764-386	Sequence 386, App	C 597	7	1.0	440	9	US-10-184-633-386	Sequence 386, App
C 525	7	1.0	440	9	US-10-198-765-386	Sequence 386, App	C 598	7	1.0	440	9	US-10-184-639-386	Sequence 386, App
C 526	7	1.0	440	9	US-10-198-768-386	Sequence 386, App	C 599	7	1.0	440	9	US-10-187-742-386	Sequence 386, App
C 527	7	1.0	440	9	US-10-198-769-386	Sequence 386, App	C 600	7	1.0	440	9	US-10-187-748-386	Sequence 386, App
C 528	7	1.0	440	9	US-10-199-305-386	Sequence 386, App	C 601	7	1.0	440	9	US-10-188-766-386	Sequence 386, App
C 529	7	1.0	440	9	US-10-199-306-386	Sequence 386, App	C 602	7	1.0	440	9	US-10-188-771-386	Sequence 386, App
C 530	7	1.0	440	9	US-10-199-310-386	Sequence 386, App	C 603	7	1.0	440	9	US-10-192-006-386	Sequence 386, App
C 531	7	1.0	440	9	US-10-199-311-386	Sequence 386, App	C 604	7	1.0	440	9	US-10-192-009-386	Sequence 386, App
C 532	7	1.0	440	9	US-10-199-314-386	Sequence 386, App	C 605	7	1.0	440	9	US-10-192-008-386	Sequence 386, App
C 533	7	1.0	440	9	US-10-199-317-386	Sequence 386, App	C 606	7	1.0	440	9	US-10-192-012-386	Sequence 386, App
C 534	7	1.0	440	9	US-10-199-665-386	Sequence 386, App	C 607	7	1.0	440	9	US-10-192-014-386	Sequence 386, App
C 535	7	1.0	440	9	US-10-199-666-386	Sequence 386, App	C 608	7	1.0	440	9	US-10-192-016-386	Sequence 386, App
C 536	7	1.0	440	9	US-10-199-669-386	Sequence 386, App	C 609	7	1.0	440	9	US-10-194-362-386	Sequence 386, App
C 537	7	1.0	440	9	US-10-201-533-386	Sequence 386, App	C 610	7	1.0	440	9	US-10-194-364-386	Sequence 386, App
C 538	7	1.0	440	9	US-10-201-770-386	Sequence 386, App	C 611	7	1.0	440	9	US-10-194-335-386	Sequence 386, App
C 539	7	1.0	440	9	US-10-201-855-386	Sequence 386, App	C 612	7	1.0	440	9	US-10-194-424-386	Sequence 386, App
C 540	7	1.0	440	9	US-10-201-856-386	Sequence 386, App	C 613	7	1.0	440	9	US-10-194-458-386	Sequence 386, App
C 541	7	1.0	440	9	US-10-202-469-386	Sequence 386, App	C 614	7	1.0	440	9	US-10-194-459-386	Sequence 386, App
C 542	7	1.0	440	9	US-10-202-470-386	Sequence 386, App	C 615	7	1.0	440	9	US-10-194-488-386	Sequence 386, App
C 543	7	1.0	440	9	US-10-202-476-386	Sequence 386, App	C 616	7	1.0	440	9	US-10-196-886-386	Sequence 386, App
C 544	7	1.0	440	9	US-10-202-934-386	Sequence 386, App	C 617	7	1.0	440	9	US-10-196-891-386	Sequence 386, App
C 545	7	1.0	440	9	US-10-202-935-386	Sequence 386, App	C 618	7	1.0	440	9	US-10-196-746-386	Sequence 386, App
C 546	7	1.0	440	9	US-10-202-936-386	Sequence 386, App	C 619	7	1.0	440	9	US-10-196-752-386	Sequence 386, App
C 547	7	1.0	440	9	US-10-202-939-386	Sequence 386, App	C 620	7	1.0	440	9	US-10-196-753-386	Sequence 386, App
C 548	7	1.0	440	9	US-10-205-504-386	Sequence 386, App	C 621	7	1.0	440	9	US-10-196-754-386	Sequence 386, App
C 549	7	1.0	440	9	US-10-205-509-386	Sequence 386, App	C 622	7	1.0	440	9	US-10-196-751-386	Sequence 386, App
C 550	7	1.0	440	9	US-10-205-895-386	Sequence 386, App	C 623	7	1.0	440	9	US-10-197-692-386	Sequence 386, App
C 551	7	1.0	440	9	US-10-205-899-386	Sequence 386, App	C 624	7	1.0	440	9	US-10-197-693-386	Sequence 386, App
C 552	7	1.0	440	9	US-10-205-900-386	Sequence 386, App	C 625	7	1.0	440	9	US-10-197-696-386	Sequence 386, App
C 553	7	1.0	440	9	US-10-205-909-386	Sequence 386, App	C 626	7	1.0	440	9	US-10-197-698-386	Sequence 386, App
C 554	7	1.0	440	9	US-10-227-693-134	Sequence 134, App	C 627	7	1.0	440	9	US-10-197-703-386	Sequence 386, App
C 555	7	1.0	440	9	US-10-183-002-386	Sequence 386, App	C 628	7	1.0	440	9	US-10-197-711-386	Sequence 386, App
C 556	7	1.0	440	9	US-10-184-638-386	Sequence 386, App	C 629	7	1.0	440	9	US-10-198-757-386	Sequence 386, App
C 557	7	1.0	440	9	US-10-184-638-386	Sequence 386, App	C 630	7	1.0	440	9	US-10-198-761-386	Sequence 386, App
C 558	7	1.0	440	9	US-10-187-752-386	Sequence 386, App	C 631	7	1.0	440	9	US-10-198-762-386	Sequence 386, App
C 559	7	1.0	440	9	US-10-187-887-386	Sequence 386, App	C 632	7	1.0	440	9	US-10-198-763-386	Sequence 386, App
C 560	7	1.0	440	9	US-10-194-461-386	Sequence 386, App	C 633	7	1.0	440	9	US-10-198-767-386	Sequence 386, App
C 561	7	1.0	440	9	US-10-195-892-386	Sequence 386, App	C 634	7	1.0	440	9	US-10-199-301-386	Sequence 386, App
C 562	7	1.0	440	9	US-10-196-751-386	Sequence 386, App	C 635	7	1.0	440	9	US-10-199-307-386	Sequence 386, App
C 563	7	1.0	440	9	US-10-197-694-386	Sequence 386, App	C 636	7	1.0	440	9	US-10-199-312-386	Sequence 386, App
C 564	7	1.0	440	9	US-10-197-697-386	Sequence 386, App	C 637	7	1.0	440	9	US-10-199-315-386	Sequence 386, App
C 565	7	1.0	440	9	US-10-197-707-386	Sequence 386, App	C 638	7	1.0	440	9	US-10-199-316-386	Sequence 386, App
C 566	7	1.0	440	9	US-10-199-303-386	Sequence 386, App	C 639	7	1.0	440	9	US-10-199-457-386	Sequence 386, App
C 567	7	1.0	440	9	US-10-199-318-386	Sequence 386, App	C 640	7	1.0	440	9	US-10-199-459-386	Sequence 386, App
C 568	7	1.0	440	9	US-10-199-458-386	Sequence 386, App	C 641	7	1.0	440	9	US-10-199-460-386	Sequence 386, App
C 569	7	1.0	440	9	US-10-199-462-386	Sequence 386, App	C 642	7	1.0	440	9	US-10-199-461-386	Sequence 386, App
C 570	7	1.0	440	9	US-10-201-324-386	Sequence 386, App	C 643	7	1.0	440	9	US-10-199-667-386	Sequence 386, App
C 571	7	1.0	440	9	US-10-201-328-386	Sequence 386, App	C 644	7	1.0	440	9	US-10-199-673-386	Sequence 386, App
C 572	7	1.0	440	9	US-10-201-527-386	Sequence 386, App	C 645	7	1.0	440	9	US-10-201-331-386	Sequence 386, App
C 573	7	1.0	440	9	US-10-201-527-386	Sequence 386, App	C 646	7	1.0	440	9	US-10-201-332-386	Sequence 386, App
C 574	7	1.0	440	9	US-10-201-529-386	Sequence 386, App	C 647	7	1.0	440	9	US-10-201-326-386	Sequence 386, App
C 575	7	1.0	440	9	US-10-201-530-386	Sequence 386, App	C 648	7	1.0	440	9	US-10-201-532-386	Sequence 386, App
C 576	7	1.0	440	9	US-10-202-408-386	Sequence 386, App	C 649	7	1.0	440	9	US-10-201-533-386	Sequence 386, App
C 577	7	1.0	440	9	US-10-202-409-386	Sequence 386, App	C 650	7	1.0	440	9	US-10-201-535-386	Sequence 386, App
C 578	7	1.0	440	9	US-10-202-411-386	Sequence 386, App	C 651	7	1.0	440	9	US-10-201-769-386	Sequence 386, App
C 579	7	1.0	440	9	US-10-202-472-386	Sequence 386, App	C 652	7	1.0	440	9	US-10-201-771-386	Sequence 386, App
C 580	7	1.0	440	9	US-10-205-502-386	Sequence 386, App	C 653	7	1.0	440	9	US-10-201-854-386	Sequence 386, App
C 581	7	1.0	440	9	US-10-205-507-386	Sequence 386, App	C 654	7	1.0	440	9	US-10-202-410-386	Sequence 386, App
C 582	7	1.0	440	9	US-10-205-511-386	Sequence 386, App	C 655	7	1.0	440	9	US-10-202-473-386	Sequence 386, App
C 583	7	1.0	440	9	US-10-205-902-386	Sequence 386, App	C 656	7	1.0	440	9	US-10-202-474-386	Sequence 386, App
C 584	7	1.0	440	9	US-10-205-907-386	Sequence 386, App	C 657	7	1.0	440	9	US-10-205-503-386	Sequence 386, App
C 585	7	1.0	440	9	US-10-194-456-386	Sequence 386, App	C 658	7	1.0	440	9	US-10-205-512-386	Sequence 386, App
C 586	7	1.0	440	9	US-10-196-758-386	Sequence 386, App	C 659	7	1.0	440	9	US-10-205-892-386	Sequence 386, App
C 587	7	1.0	440	9	US-10-198-770-386	Sequence 386, App	C 660	7	1.0	440	9	US-10-205-894-386	Sequence 386, App
C 588	7	1.0	440	9	US-10-199-308-386	Sequence 386, App	C 661	7	1.0	440	9	US-10-205-896-386	Sequence 386, App

C 662	7	1.0	440	9	US-10-205-898-386	Sequence 386, App	735	7	1.0	448	10	US-09-815-242-12574	Sequence 12574, A
C 663	7	1.0	440	9	US-10-205-901-386	Sequence 386, App	C 736	7	1.0	449	9	US-09-990-046-29	Sequence 29, App1
C 664	7	1.0	440	9	US-10-205-903-386	Sequence 386, App	C 737	7	1.0	449	9	US-09-733-634-30	Sequence 30, App1
C 665	7	1.0	440	9	US-10-206-909-386	Sequence 386, App	C 738	7	1.0	449	10	US-09-764-864-1163	Sequence 1163, Ap
C 666	7	1.0	440	9	US-10-206-910-386	Sequence 386, App	C 739	7	1.0	442	10	US-09-815-242-10893	Sequence 10893, A
C 667	7	1.0	440	9	US-10-206-911-386	Sequence 386, App	C 740	7	1.0	453	10	US-09-815-242-10872	Sequence 10872, A
C 668	7	1.0	440	9	US-10-206-912-386	Sequence 386, App	C 741	7	1.0	444	10	US-09-815-242-10872	Sequence 10872, A
C 669	7	1.0	440	9	US-10-206-913-386	Sequence 386, App	C 742	7	1.0	456	9	US-09-764-864-985	Sequence 985, App
C 670	7	1.0	440	9	US-10-206-914-386	Sequence 386, App	C 743	7	1.0	458	10	US-09-739-457-4	Sequence 4, App1
C 671	7	1.0	440	9	US-10-206-920-386	Sequence 386, App	C 744	7	1.0	459	9	US-09-769-783-85	Sequence 85, App1
C 672	7	1.0	440	9	US-10-206-921-386	Sequence 386, App	C 745	7	1.0	460	9	US-09-738-626-6145	Sequence 6145, Ap
C 673	7	1.0	440	9	US-10-206-923-386	Sequence 386, App	C 746	7	1.0	464	9	US-09-988-462-22	Sequence 22, App1
C 674	7	1.0	440	9	US-10-206-925-386	Sequence 386, App	C 747	7	1.0	459	10	US-09-925-302-627	Sequence 627, App
C 675	7	1.0	440	9	US-10-206-926-386	Sequence 386, App	C 748	7	1.0	480	9	US-10-084-418-8	Sequence 8, App1
C 676	7	1.0	440	9	US-10-206-927-386	Sequence 386, App	C 749	7	1.0	480	9	US-09-736-457-336	Sequence 336, App
C 677	7	1.0	440	9	US-10-207-916-386	Sequence 386, App	C 750	7	1.0	480	9	US-09-902-941-336	Sequence 336, App
C 678	7	1.0	440	9	US-10-207-917-386	Sequence 386, App	C 751	7	1.0	480	9	US-09-849-626-336	Sequence 336, App
C 679	7	1.0	440	9	US-10-207-918-386	Sequence 386, App	C 752	7	1.0	480	9	US-10-017-754-336	Sequence 336, App
C 680	7	1.0	440	9	US-10-207-919-386	Sequence 386, App	C 753	7	1.0	481	10	US-09-764-864-979	Sequence 979, App
C 681	7	1.0	440	9	US-10-207-920-386	Sequence 386, App	C 754	7	1.0	482	9	US-09-746-783-66	Sequence 66, App1
C 682	7	1.0	440	9	US-10-207-925-386	Sequence 386, App	C 755	7	1.0	483	10	US-09-799-777-69	Sequence 69, App1
C 683	7	1.0	440	9	US-10-208-021-386	Sequence 386, App	C 756	7	1.0	484	10	US-09-815-242-10809	Sequence 10809, A
C 684	7	1.0	440	9	US-10-208-022-386	Sequence 386, App	C 757	7	1.0	466	9	US-10-055-475-12	Sequence 12, App1
C 685	7	1.0	440	9	US-10-208-023-386	Sequence 386, App	C 758	7	1.0	480	10	US-09-739-457-3	Sequence 3, App1
C 686	7	1.0	440	9	US-10-208-026-386	Sequence 386, App	C 759	7	1.0	480	10	US-09-739-457-6	Sequence 6, App1
C 687	7	1.0	440	9	US-10-208-029-386	Sequence 386, App	C 760	7	1.0	490	10	US-09-739-457-7	Sequence 7, App1
C 688	7	1.0	440	9	US-10-208-030-386	Sequence 386, App	C 761	7	1.0	492	9	US-09-852-797-72	Sequence 72, App1
C 689	7	1.0	440	9	US-10-232-232-386	Sequence 386, App	C 762	7	1.0	492	10	US-09-853-161-72	Sequence 72, App1
C 690	7	1.0	440	9	US-10-063-538-134	Sequence 134, App	C 763	7	1.0	482	10	US-09-852-659A-72	Sequence 72, App1
C 691	7	1.0	440	9	US-10-173-693-386	Sequence 386, App	C 764	7	1.0	501	9	US-10-051-809-38	Sequence 38, App1
C 692	7	1.0	440	9	US-10-174-578-386	Sequence 386, App	C 765	7	1.0	504	10	US-09-883-797-6	Sequence 6, App1
C 693	7	1.0	440	9	US-10-175-741-386	Sequence 386, App	C 766	7	1.0	513	10	US-09-818-264-4	Sequence 4, App1
C 694	7	1.0	440	9	US-10-175-750-386	Sequence 386, App	C 767	7	1.0	518	9	US-09-976-297-2	Sequence 2, App1
C 695	7	1.0	440	9	US-10-176-986-386	Sequence 386, App	C 768	7	1.0	519	10	US-09-925-300-1680	Sequence 1680, Ap
C 696	7	1.0	440	9	US-10-184-641-386	Sequence 386, App	C 769	7	1.0	532	9	US-10-045-615-8	Sequence 8, App1
C 697	7	1.0	440	9	US-10-187-888-386	Sequence 386, App	C 770	7	1.0	539	9	US-10-055-475-16	Sequence 16, App1
C 698	7	1.0	440	9	US-10-194-360-386	Sequence 386, App	C 771	7	1.0	549	10	US-09-828-313-39	Sequence 39, App1
C 699	7	1.0	440	9	US-10-194-365-386	Sequence 386, App	C 772	7	1.0	550	9	US-10-095-139-16	Sequence 16, App1
C 700	7	1.0	440	9	US-10-195-895-386	Sequence 386, App	C 773	7	1.0	550	9	US-10-339-513-2	Sequence 2, App1
C 701	7	1.0	440	9	US-10-195-898-386	Sequence 386, App	C 774	7	1.0	553	9	US-09-282-862-7	Sequence 7, App1
C 702	7	1.0	440	9	US-10-196-759-386	Sequence 386, App	C 775	7	1.0	553	9	US-10-233-873A-7	Sequence 7, App1
C 703	7	1.0	440	9	US-10-199-302-386	Sequence 386, App	C 776	7	1.0	553	10	US-09-746-359A-11	Sequence 11, App1
C 704	7	1.0	440	9	US-10-201-323-386	Sequence 386, App	C 777	7	1.0	553	10	US-09-949-19-107	Sequence 1070, Ap
C 705	7	1.0	440	9	US-10-205-510-386	Sequence 386, App	C 778	7	1.0	553	10	US-09-764-864-1068	Sequence 1068, Ap
C 706	7	1.0	440	9	US-10-205-891-386	Sequence 386, App	C 779	7	1.0	571	10	US-09-764-864-1068	Sequence 1068, Ap
C 707	7	1.0	440	9	US-10-205-904-386	Sequence 386, App	C 780	7	1.0	579	10	US-09-815-242-15821	Sequence 5821, Ap
C 708	7	1.0	440	9	US-10-206-917-386	Sequence 386, App	C 781	7	1.0	579	10	US-09-815-242-12943	Sequence 12943, A
C 709	7	1.0	440	9	US-10-207-923-386	Sequence 386, App	C 782	7	1.0	584	9	US-10-050-786-7	Sequence 7, App1
C 710	7	1.0	440	9	US-10-207-924-386	Sequence 386, App	C 783	7	1.0	585	9	US-10-173-519-4	Sequence 4, App1
C 711	7	1.0	440	9	US-10-208-028-386	Sequence 386, App	C 784	7	1.0	587	9	US-10-173-519-5	Sequence 5, App1
C 712	7	1.0	440	9	US-10-063-559-134	Sequence 134, App	C 785	7	1.0	587	9	US-10-167-994-12	Sequence 12, App1
C 713	7	1.0	440	9	US-10-121-062-386	Sequence 386, App	C 786	7	1.0	594	9	US-10-028-072-10	Sequence 10, App1
C 714	7	1.0	440	9	US-10-175-753-386	Sequence 386, App	C 787	7	1.0	594	9	US-10-121-049-10	Sequence 10, App1
C 715	7	1.0	440	9	US-10-180-553-386	Sequence 386, App	C 788	7	1.0	594	9	US-10-123-904-10	Sequence 10, App1
C 716	7	1.0	440	9	US-10-201-327-386	Sequence 386, App	C 789	7	1.0	594	9	US-10-140-470-10	Sequence 10, App1
C 717	7	1.0	440	9	US-10-063-595-134	Sequence 134, App	C 790	7	1.0	594	9	US-10-175-746-10	Sequence 10, App1
C 718	7	1.0	440	9	US-10-173-696-386	Sequence 386, App	C 791	7	1.0	594	9	US-10-176-918-10	Sequence 10, App1
C 719	7	1.0	440	9	US-10-183-003-386	Sequence 386, App	C 792	7	1.0	594	9	US-10-176-921-10	Sequence 10, App1
C 720	7	1.0	440	9	US-10-183-016-386	Sequence 386, App	C 793	7	1.0	594	9	US-10-137-865-10	Sequence 10, App1
C 721	7	1.0	440	9	US-10-125-923A-386	Sequence 386, App	C 794	7	1.0	594	9	US-10-140-474-10	Sequence 10, App1
C 722	7	1.0	440	9	US-10-176-491-386	Sequence 386, App	C 795	7	1.0	594	9	US-10-142-431-10	Sequence 10, App1
C 723	7	1.0	440	9	US-10-176-979-386	Sequence 386, App	C 796	7	1.0	594	9	US-10-143-114-10	Sequence 10, App1
C 724	7	1.0	440	9	US-10-187-592-386	Sequence 386, App	C 797	7	1.0	594	9	US-10-140-002-10	Sequence 10, App1
C 725	7	1.0	440	9	US-10-197-691-386	Sequence 386, App	C 798	7	1.0	594	9	US-10-142-419-10	Sequence 10, App1
C 726	7	1.0	440	9	US-10-196-771-386	Sequence 386, App	C 799	7	1.0	594	9	US-10-123-262-10	Sequence 10, App1
C 727	7	1.0	440	9	US-10-174-575A-386	Sequence 386, App	C 800	7	1.0	594	9	US-10-142-423-10	Sequence 10, App1
C 728	7	1.0	440	9	US-10-179-520-386	Sequence 386, App	C 801	7	1.0	594	9	US-10-121-050-10	Sequence 10, App1
C 729	7	1.0	440	9	US-10-201-325-386	Sequence 386, App	C 802	7	1.0	594	9	US-10-141-755-10	Sequence 10, App1
C 730	7	1.0	440	9	US-10-202-941-386	Sequence 386, App	C 803	7	1.0	594	9	US-10-143-032-10	Sequence 10, App1
C 731	7	1.0	440	9	US-10-205-910-386	Sequence 386, App	C 804	7	1.0	594	9	US-10-123-108-10	Sequence 10, App1
C 732	7	1.0	440	9	US-10-179-526-386	Sequence 386, App	C 805	7	1.0	594	9	US-10-123-236-10	Sequence 10, App1
C 733	7	1.0	440	12	US-10-006-867-134	Sequence 134, App	C 806	7	1.0	594	9	US-10-123-261-10	Sequence 10, App1
C 734	7	1.0	440	12	US-10-052-586-386	Sequence 386, App	C 807	7	1.0	594	9	US-10-140-921-10	Sequence 10, App1

808	7	1.0	594	9	US-10-140-928-10	Sequence 10, Appl	881	7	1.0	594	9	US-10-147-484-10	Sequence 10, Appl
809	7	1.0	594	9	US-10-121-045-10	Sequence 10, Appl	882	7	1.0	594	9	US-10-147-492-10	Sequence 10, Appl
810	7	1.0	594	9	US-10-123-292-10	Sequence 10, Appl	883	7	1.0	594	9	US-10-147-508-10	Sequence 10, Appl
811	7	1.0	594	9	US-10-123-903-10	Sequence 10, Appl	884	7	1.0	594	9	US-10-147-512-10	Sequence 10, Appl
812	7	1.0	594	9	US-10-124-819-10	Sequence 10, Appl	885	7	1.0	594	9	US-10-158-782-10	Sequence 10, Appl
813	7	1.0	594	9	US-10-124-822-10	Sequence 10, Appl	886	7	1.0	594	9	US-10-175-735-10	Sequence 10, Appl
814	7	1.0	594	9	US-10-140-925-10	Sequence 10, Appl	887	7	1.0	594	9	US-10-123-905-10	Sequence 10, Appl
815	7	1.0	594	9	US-10-160-498-10	Sequence 10, Appl	888	7	1.0	594	9	US-10-123-907-10	Sequence 10, Appl
816	7	1.0	594	9	US-10-121-041-10	Sequence 10, Appl	889	7	1.0	594	9	US-10-124-815-10	Sequence 10, Appl
817	7	1.0	594	9	US-10-121-043-10	Sequence 10, Appl	890	7	1.0	594	9	US-10-125-921A-10	Sequence 10, Appl
818	7	1.0	594	9	US-10-121-047-10	Sequence 10, Appl	891	7	1.0	594	9	US-10-125-928A-10	Sequence 10, Appl
819	7	1.0	594	9	US-10-123-215-10	Sequence 10, Appl	892	7	1.0	594	9	US-10-127-821A-10	Sequence 10, Appl
820	7	1.0	594	9	US-10-123-902-10	Sequence 10, Appl	893	7	1.0	594	9	US-10-127-822A-10	Sequence 10, Appl
821	7	1.0	594	9	US-10-123-908-10	Sequence 10, Appl	894	7	1.0	594	9	US-10-127-824A-10	Sequence 10, Appl
822	7	1.0	594	9	US-10-123-909-10	Sequence 10, Appl	895	7	1.0	594	9	US-10-127-826A-10	Sequence 10, Appl
823	7	1.0	594	9	US-10-123-910-10	Sequence 10, Appl	896	7	1.0	594	9	US-10-127-834A-10	Sequence 10, Appl
824	7	1.0	594	9	US-10-124-813-10	Sequence 10, Appl	897	7	1.0	594	9	US-10-127-836A-10	Sequence 10, Appl
825	7	1.0	594	9	US-10-124-817-10	Sequence 10, Appl	898	7	1.0	594	9	US-10-127-828A-10	Sequence 10, Appl
826	7	1.0	594	9	US-10-124-824-10	Sequence 10, Appl	899	7	1.0	594	9	US-10-127-830A-10	Sequence 10, Appl
827	7	1.0	594	9	US-10-125-922-10	Sequence 10, Appl	900	7	1.0	594	9	US-10-127-832A-10	Sequence 10, Appl
828	7	1.0	594	9	US-10-125-924-10	Sequence 10, Appl	901	7	1.0	594	9	US-10-127-833A-10	Sequence 10, Appl
829	7	1.0	594	9	US-10-127-825A-10	Sequence 10, Appl	902	7	1.0	594	9	US-10-127-834A-10	Sequence 10, Appl
830	7	1.0	594	9	US-10-127-829A-10	Sequence 10, Appl	903	7	1.0	594	9	US-10-127-836A-10	Sequence 10, Appl
831	7	1.0	594	9	US-10-127-835A-10	Sequence 10, Appl	904	7	1.0	594	9	US-10-127-841A-10	Sequence 10, Appl
832	7	1.0	594	9	US-10-127-839A-10	Sequence 10, Appl	905	7	1.0	594	9	US-10-127-844A-10	Sequence 10, Appl
833	7	1.0	594	9	US-10-127-901A-10	Sequence 10, Appl	906	7	1.0	594	9	US-10-128-687A-10	Sequence 10, Appl
834	7	1.0	594	9	US-10-128-693A-10	Sequence 10, Appl	907	7	1.0	594	9	US-10-128-688A-10	Sequence 10, Appl
835	7	1.0	594	9	US-10-131-813A-10	Sequence 10, Appl	908	7	1.0	594	9	US-10-128-689A-10	Sequence 10, Appl
836	7	1.0	594	9	US-10-131-818A-10	Sequence 10, Appl	909	7	1.0	594	9	US-10-128-694A-10	Sequence 10, Appl
837	7	1.0	594	9	US-10-131-823A-10	Sequence 10, Appl	910	7	1.0	594	9	US-10-131-825A-10	Sequence 10, Appl
838	7	1.0	594	9	US-10-131-824A-10	Sequence 10, Appl	911	7	1.0	594	9	US-10-131-826A-10	Sequence 10, Appl
839	7	1.0	594	9	US-10-131-824A-10	Sequence 10, Appl	912	7	1.0	594	9	US-10-131-827A-10	Sequence 10, Appl
840	7	1.0	594	9	US-10-131-830A-10	Sequence 10, Appl	913	7	1.0	594	9	US-10-131-817A-10	Sequence 10, Appl
841	7	1.0	594	9	US-10-131-837A-10	Sequence 10, Appl	914	7	1.0	594	9	US-10-131-821A-10	Sequence 10, Appl
842	7	1.0	594	9	US-10-140-860-10	Sequence 10, Appl	915	7	1.0	594	9	US-10-131-822A-10	Sequence 10, Appl
843	7	1.0	594	9	US-10-142-417-10	Sequence 10, Appl	916	7	1.0	594	9	US-10-131-828A-10	Sequence 10, Appl
844	7	1.0	594	9	US-10-147-500-10	Sequence 10, Appl	917	7	1.0	594	9	US-10-131-835A-10	Sequence 10, Appl
845	7	1.0	594	9	US-10-147-502-10	Sequence 10, Appl	918	7	1.0	594	9	US-10-137-864A-10	Sequence 10, Appl
846	7	1.0	594	9	US-10-147-515-10	Sequence 10, Appl	919	7	1.0	594	9	US-10-137-869A-10	Sequence 10, Appl
847	7	1.0	594	9	US-10-147-517-10	Sequence 10, Appl	920	7	1.0	594	9	US-10-147-523-10	Sequence 10, Appl
848	7	1.0	594	9	US-10-147-519-10	Sequence 10, Appl	921	7	1.0	594	9	US-10-158-785-10	Sequence 10, Appl
849	7	1.0	594	9	US-10-147-526-10	Sequence 10, Appl	922	7	1.0	594	9	US-10-121-042-10	Sequence 10, Appl
850	7	1.0	594	9	US-10-147-527-10	Sequence 10, Appl	923	7	1.0	594	9	US-10-123-912-10	Sequence 10, Appl
851	7	1.0	594	9	US-10-152-395-10	Sequence 10, Appl	924	7	1.0	612	9	US-09-875-353-4	Sequence 4, Appl1
852	7	1.0	594	9	US-10-157-782-10	Sequence 10, Appl	925	7	1.0	623	9	US-09-815-379-2	Sequence 2, Appl1
853	7	1.0	594	9	US-10-121-040-10	Sequence 10, Appl	926	7	1.0	633	9	US-10-086-464-11	Sequence 11, Appl
854	7	1.0	594	9	US-10-121-056-10	Sequence 10, Appl	927	7	1.0	635	9	US-10-214-932-86	Sequence 86, Appl
855	7	1.0	594	9	US-10-121-061-10	Sequence 10, Appl	928	7	1.0	637	9	US-09-948-137A-4	Sequence 4, Appl1
856	7	1.0	594	9	US-10-123-235-10	Sequence 10, Appl	929	7	1.0	652	10	US-09-789-919-96	Sequence 96, Appl
857	7	1.0	594	9	US-10-124-818-10	Sequence 10, Appl	930	7	1.0	653	9	US-10-023-437-67	Sequence 67, Appl
858	7	1.0	594	9	US-10-125-926A-10	Sequence 10, Appl	931	7	1.0	653	9	US-09-820-843A-26	Sequence 26, Appl
859	7	1.0	594	9	US-10-125-930A-10	Sequence 10, Appl	932	7	1.0	695	9	US-09-738-626-5527	Sequence 5527, Ap
860	7	1.0	594	9	US-10-127-831A-10	Sequence 10, Appl	933	7	1.0	695	10	US-09-810-796-15	Sequence 15, Appl
861	7	1.0	594	9	US-10-127-837A-10	Sequence 10, Appl	934	7	1.0	696	10	US-09-813-148-6	Sequence 6, Appl1
862	7	1.0	594	9	US-10-127-838B-10	Sequence 10, Appl	935	7	1.0	704	10	US-09-815-242-11805	Sequence 11805, A
863	7	1.0	594	9	US-10-127-842A-10	Sequence 10, Appl	936	7	1.0	705	9	US-10-043-487-269	Sequence 269, App
864	7	1.0	594	9	US-10-127-843A-10	Sequence 10, Appl	937	7	1.0	707	9	US-10-068-674-25	Sequence 25, Appl
865	7	1.0	594	9	US-10-127-845A-10	Sequence 10, Appl	938	7	1.0	722	9	US-10-128-870-23	Sequence 23, Appl
866	7	1.0	594	9	US-10-127-846A-10	Sequence 10, Appl	939	7	1.0	722	9	US-10-131-685-23	Sequence 23, Appl
867	7	1.0	594	9	US-10-127-848A-10	Sequence 10, Appl	940	7	1.0	724	9	US-10-108-605-507	Sequence 307, App
868	7	1.0	594	9	US-10-127-849A-10	Sequence 10, Appl	941	7	1.0	747	9	US-10-045-815-6	Sequence 6, Appl1
869	7	1.0	594	9	US-10-127-850A-10	Sequence 10, Appl	942	7	1.0	763	9	US-09-989-339-4	Sequence 4, Appl
870	7	1.0	594	9	US-10-127-851A-10	Sequence 10, Appl	943	7	1.0	763	9	US-09-769-787-63	Sequence 63, Appl
871	7	1.0	594	9	US-10-128-684A-10	Sequence 10, Appl	944	7	1.0	765	12	US-10-044-303-1	Sequence 1, Appl1
872	7	1.0	594	9	US-10-128-686A-10	Sequence 10, Appl	945	7	1.0	766	10	US-09-815-242-10103	Sequence 10103, A
873	7	1.0	594	9	US-10-128-690A-10	Sequence 10, Appl	946	7	1.0	808	9	US-10-108-605-127	Sequence 127, App
874	7	1.0	594	9	US-10-128-691A-10	Sequence 10, Appl	947	7	1.0	815	10	US-09-371-307-88	Sequence 88, Appl
875	7	1.0	594	9	US-10-131-819A-10	Sequence 10, Appl	948	7	1.0	816	9	US-09-796-7208-2	Sequence 2, Appl1
876	7	1.0	594	9	US-10-131-829A-10	Sequence 10, Appl	949	7	1.0	819	9	US-09-796-7208-4	Sequence 4, Appl1
877	7	1.0	594	9	US-10-131-836A-10	Sequence 10, Appl	950	7	1.0	823	10	US-09-815-242-11348	Sequence 1348, A
878	7	1.0	594	9	US-10-137-868-10	Sequence 10, Appl	951	7	1.0	840	10	US-09-815-242-5198	Sequence 5198, Ap
879	7	1.0	594	9	US-10-146-729-10	Sequence 10, Appl	952	7	1.0	844	10	US-09-813-148-4	Sequence 4, Appl
880	7	1.0	594	9	US-10-146-791-10	Sequence 10, Appl	953	7	1.0	844	10	US-09-810-796-14	Sequence 14, Appl

```
C 954 7 1.0 871 9 US-10-128-870-20 Sequence 20, Appl
C 955 7 1.0 871 9 US-10-131-685-20 Sequence 20, Appl
C 956 7 1.0 878 10 US-10-169-048-42 Sequence 42, Appl
C 957 7 1.0 878 10 US-09-912-020-364 Sequence 364, Appl
C 958 7 1.0 881 10 US-09-982-736-2 Sequence 2, Appl
C 959 7 1.0 888 10 US-09-810-796-5 Sequence 5, Appl
C 960 7 1.0 890 9 US-09-373-658-4 Sequence 4, Appl
C 961 7 1.0 897 10 US-09-810-796-4 Sequence 4, Appl
C 962 7 1.0 911 12 US-10-141-132-2 Sequence 2, Appl
C 963 7 1.0 922 10 US-09-908-1538-42 Sequence 42, Appl
C 964 7 1.0 923 10 US-09-825-147-2 Sequence 2, Appl
C 965 7 1.0 926 10 US-09-816-685-2 Sequence 2, Appl
C 966 7 1.0 932 10 US-09-813-148-2 Sequence 2, Appl
C 967 7 1.0 966 10 US-09-771-161A-197 Sequence 197, Appl
C 968 7 1.0 982 10 US-09-991-496-95 Sequence 95, Appl
C 969 7 1.0 982 10 US-09-841-132-176 Sequence 176, Appl
C 970 7 1.0 982 10 US-09-874-823-95 Sequence 95, Appl
C 971 7 1.0 1002 9 US-09-988-117-3 Sequence 3, Appl
C 972 7 1.0 1002 10 US-09-812-471-3 Sequence 3, Appl
C 973 7 1.0 1002 10 US-09-812-633-3 Sequence 3, Appl
C 974 7 1.0 1006 10 US-09-841-132-190 Sequence 190, Appl
C 975 7 1.0 1014 9 US-10-054-044A-4 Sequence 4, Appl
C 976 7 1.0 1014 9 US-10-120-604-4 Sequence 4, Appl
C 977 7 1.0 1014 9 US-10-120-604-139 Sequence 139, Appl
C 978 7 1.0 1016 12 US-10-044-303-2 Sequence 2, Appl
C 979 7 1.0 1021 9 US-10-101-464A-954 Sequence 954, Appl
C 980 7 1.0 1164 10 US-09-950-046A-2 Sequence 2, Appl
C 981 7 1.0 1242 9 US-09-925-299-911 Sequence 911, Appl
C 982 7 1.0 1242 10 US-09-925-299-911 Sequence 911, Appl
C 983 7 1.0 1300 10 US-09-815-242-4903 Sequence 4903, Appl
C 984 7 1.0 1300 10 US-09-815-242-10906 Sequence 10906, A
C 985 7 1.0 1306 9 US-10-017-724-4 Sequence 4, Appl
C 986 7 1.0 1311 1 US-08-954-701A-4 Sequence 4, Appl
C 987 7 1.0 1338 9 US-10-029-413A-10 Sequence 10, Appl
C 988 7 1.0 1356 9 US-10-090-280-33 Sequence 33, Appl
C 989 7 1.0 1359 9 US-10-090-280-34 Sequence 34, Appl
C 990 7 1.0 1375 10 US-09-740-274-4 Sequence 4, Appl
C 991 7 1.0 1403 9 US-10-123-036-4 Sequence 4, Appl
C 992 7 1.0 1403 9 US-10-108-605-93 Sequence 93, Appl
C 993 7 1.0 1518 10 US-09-801-368-152 Sequence 152, Appl
C 994 7 1.0 1620 10 US-09-827-949-2 Sequence 2, Appl
C 995 7 1.0 1641 9 US-09-991-496-96 Sequence 96, Appl
C 996 7 1.0 1641 10 US-09-874-923-96 Sequence 96, Appl
C 997 7 1.0 1643 9 US-09-738-626-6894 Sequence 6894, Appl
C 998 7 1.0 1647 9 US-09-824-574-4 Sequence 4, Appl
C 999 7 1.0 1663 9 US-09-843-758-41 Sequence 41, Appl
C1000 7 1.0 1663 10 US-09-875-519A-22 Sequence 22, Appl
```

## ALIGNMENTS

```
RESULT 1
US-09-764-891-3857
; Sequence 3857, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764, 891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3857
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3857
```

```
Alignment Scores:
Pred. No.: 3,15e-68 Length: 106
Score: 84.00 Matches: 84
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.60% Indels: 0
DB: 9 Gaps: 0
```

US-09-714-865-15 (1-2172) x US-09-764-891-3857 (1-106)

OY 1885 GGGCGTCTGCTGCTGTTGGGAATCTGCGACAGCATTTTCCTTTTGTGATCTTGATCG 1944

DB 11 G1YArghThrlYArYgCYsG1YAsnThrG1YArGAlAlLeSerPhePheAspLeuG1Ser 30

OY 1945 GATPACCATTTAGACACAGCCTCTAGTAAAGTTTGCACATGCTCAACAGATGTTCT 2004

DB 31 Aspshh1sLeuAlaGlnProLeuValValLeuThrAspAlaGlnGlnAspValPro 50

OY 2005 GCATGTTGGAGAAATGGCTTTAGTACATACATTCCTGCTTCACTGAGTATACAGA 2064

DB 51 AlatrPleuGlnGlnIleAlaPheSerThrTyrlleProGlyPheSerGlySerThrArg 70

OY 2065 GGAACGTGTTTGCATGATGATGATACCCAGAAAGGCGCAAGCACTTGAACACAGCTGGG 2124

DB 71 G1YAsnValPheAlaSerValAspThrArgG1YSerThrLeuAsnThrAlaGly 90

OY 2125 TTTTCTCTTCA 2136

DB 91 PheSerSerSer 94

RESULT 2

US-10-195-117-3

; Sequence 3, Application US/10195117

; Publication No. US20030092083A1

; GENERAL INFORMATION:

; APPLICANT: Inogen Co., Ltd.

; APPLICANT: Jeoung, Doo-Il

; APPLICANT: Cho, Bomsoo

; APPLICANT: Lim, Yoon

; APPLICANT: Park, Saeyoung

; APPLICANT: Lee, Daeyeon

; APPLICANT: Bang, Yung-Jue

; APPLICANT: Yang, Hankwang

; APPLICANT: Kim, Dae-Ke

; TITLE OF INVENTION: CAGE Antigen

; FILE REFERENCE: S8258-00002

; CURRENT APPLICATION NUMBER: US/10/195,117

; CURRENT FILING DATE: 2002-07-11

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 238

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-195-117-3

; Alignment Scores:

Pred. No.: 6,54e-05 Length: 238

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.07% Indels: 0

DB: 9 Gaps: 0

US-09-714-865-15 (1-2172) x US-10-195-117-3 (1-238)

OY 1324 TACTTAGTTTGGATGAAGCTGATCCGATGTTGGATATGGATTT 1368

DB 153 TyrIeuValIeuAspGlnAlaAspArgMetLeuAspMetGlyPhe 167

RESULT 3

US-10-195-117-4

Sequence 4, Application US/10195117  
Publication No. US20030092083A1  
GENERAL INFORMATION:  
APPLICANT: Inogen Co., Ltd.  
APPLICANT: Jeonung, Doo-Il  
APPLICANT: Cho, Bomsoo  
APPLICANT: Lim, Yoon  
APPLICANT: Park, Saeyoung  
APPLICANT: Lee, Daeyeon  
APPLICANT: Bang, Yung-Jue  
APPLICANT: Yang, Hankwang  
APPLICANT: Kim, Dae-Kee  
TITLE OF INVENTION: CAGE Antigen  
FILE REFERENCE: 59258-00002  
CURRENT APPLICATION NUMBER: US/10/195,117  
CURRENT FILING DATE: 2002-07-11  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 238  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-195-117-4  
Alignment Scores:  
Pred. No.: 6,54e-05 Length: 238  
Score: 15.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.07% Indels: 0  
DB: Gaps: 0  
US-09-714-865-15 (1-2172) x US-10-195-117-4 (1-238)  
QY 1324 TACTTACTTTGGATGAAGCTGATCGCATGTGGATATGGGTTT 1368  
DB 153 TyrluValleuaspGluAlaaspArgMetleuaspMetGlyPhe 167  
RESULT 4  
US-09-925-301-1439  
Sequence 1439, Application US/09925301  
Patent No. US20020052308A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA106  
CURRENT APPLICATION NUMBER: US/09/925,301  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05882  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
NUMBER OF SEQ ID NOS: 1694  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1439  
LENGTH: 343  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (244)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (305)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (325)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (328)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (340)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-301-1439  
Alignment Scores:  
Pred. No.: 6.17e-05 Length: 343  
Score: 15.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.07% Indels: 0  
DB: Gaps: 0  
US-09-714-865-15 (1-2172) x US-09-925-301-1439 (1-343)  
QY 1324 TACTTACTTTGGATGAAGCTGATCGCATGTGGATATGGGTTT 1368  
DB 153 TyrluValleuaspGluAlaaspArgMetleuaspMetGlyPhe 167  
RESULT 5  
US-10-228-897-13  
Sequence 13, Application US/10228897  
Publication No. US20030092043A1  
GENERAL INFORMATION:  
APPLICANT: Fisher, Paul B.  
APPLICANT: Kang, Dong-Chul  
APPLICANT: GopalKrishnan, Rahul V.  
TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED  
FILE REFERENCE: A34614-A-PCT-USA (070050,2121)  
CURRENT APPLICATION NUMBER: US/10/228,897  
CURRENT FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: PCT/US01/06960  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 09/515,363  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 500  
TYPE: PRT  
ORGANISM: molgula oculata  
US-10-228-897-13  
Alignment Scores:  
Pred. No.: 5.81e-05 Length: 500  
Score: 15.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.07% Indels: 0  
DB: Gaps: 0  
US-09-714-865-15 (1-2172) x US-10-228-897-13 (1-500)  
QY 1324 TACTTACTTTGGATGAAGCTGATCGCATGTGGATATGGGTTT 1368  
DB 244 TyrluValleuaspGluAlaaspArgMetleuaspMetGlyPhe 258  
RESULT 6  
US-10-108-605-55  
Sequence 55, Application US/10108605  
Patent No. US20020160934A1  
GENERAL INFORMATION:  
APPLICANT: Brodus, Julie  
APPLICANT: Stam, Lynn  
APPLICANT: Bachmann, Jane  
APPLICANT: Kamdar, Kim  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE  
FILE REFERENCE: 31133B  
CURRENT APPLICATION NUMBER: US/10/108,605  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: US 09/761,142  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/176,418





Query Match: 1.52% Indels: 0  
DB: 10 Gaps: 0

US-09-714-865-15 (1-2172) x US-09-815-242-10990 (1-613)

QY 994 GCTCAACAGGCTGTGGAGACTGCGGCTTT 1026

DB 49 AAGlnthrglSerglyLysthrAlaAlaPhe 59

## RESULT 9

US-09-815-242-10311  
; Sequence 10311, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10311  
; LENGTH: 646  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10311

## Alignment Scores:

Pred. No.: 0.263 Length: 646  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.52% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-815-242-10311 (1-646)

QY 994 GCTCAACAGGCTGTGGAGACTGCGGCTTT 1026

DB 67 AAGlnthrglSerglyLysthrAlaAlaPhe 77

## RESULT 10

US-09-815-242-13769  
; Sequence 13769, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR FILING DATE: 2001-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13769

LENGTH: 646

TYPE: PRT

ORGANISM: Salmonella typhi

US-09-815-242-13769

## Alignment Scores:

Pred. No.: 0.263 Length: 646  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.52% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-815-242-13769 (1-646)

QY 994 GCTCAACAGGCTGTGGAGACTGCGGCTTT 1026

DB 67 AAGlnthrglSerglyLysthrAlaAlaPhe 77

## RESULT 11

US-09-923-831-43  
; Sequence 43, Application US/09923831  
; Patent No. US20020115142A1  
; GENERAL INFORMATION:  
; APPLICANT: Martelange, Val, rie  
; APPLICANT: De Smet, Charles  
; APPLICANT: Boon-Falieu, Thierry

TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR

FILE REFERENCE: 10461/7054

CURRENT APPLICATION NUMBER: US/09/923,831

CURRENT FILING DATE: 2001-08-07

PRIOR FILING DATE: 2001-10-30

NUMBER OF SEQ ID NOS: 43

SEQ ID NO 43

LENGTH: 648

TYPE: PRT

ORGANISM: H. sapiens

US-09-923-831-43

## Alignment Scores:

Pred. No.: 0.263 Length: 648  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.52% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-923-831-43 (1-648)  
QY 1867 GAATATGTCATGAATGGCGTACTGTCGT 1899  
DB 570 GTyrvAlhIsArG1leGlyArGThrGlyArG 580

RESULT 12  
US-09-815-242-11703  
; Sequence 11703, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haseibeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zysekund, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11703  
; LENGTH: 421  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-815-242-11703

Alignment Scores:  
Pred. No.: 2.34 Length: 421  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.38% Indels: 0  
DB: 10 Gaps: 0

US-09-714-865-15 (1-2172) x US-09-815-242-11703 (1-421)  
QY 1870 TATGTCATGAATGGCGTACTGTCGT 1899  
DB 340 TyrvAlhIsArG1leGlyArGThrGlyArG 349

RESULT 13  
US-09-738-626-4359  
; Sequence 4359, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4359  
LENGTH: 424  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4359

Alignment Scores:  
Pred. No.: 2.33 Length: 424  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.38% Indels: 0  
DB: 9 Gaps: 0

US-09-714-865-15 (1-2172) x US-09-738-626-4359 (1-424)  
QY 1870 TATGTCATGAATGGCGTACTGTCGT 1899  
DB 338 TyrvAlhIsArG1leGlyArGThrGlyArG 347

RESULT 14  
US-09-971-536-52  
; Sequence 52, Application US/0971536  
; Patent No. US20020159976A1  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Hayukkala, Ilkka  
; APPLICANT: Bloksberg, Leonard  
; APPLICANT: Lubbers, Mark  
; APPLICANT: Dekker, James  
; APPLICANT: Christensson, Anna  
; APPLICANT: Holland, Rose  
; APPLICANT: O'Toole, Paul  
; APPLICANT: Reid, Julian  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods  
; FILE REFERENCE: 1043c2  
; CURRENT APPLICATION NUMBER: US/09/971,536  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160  
; PRIOR FILING DATE: 2001-08-08  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 52  
; LENGTH: 495  
; TYPE: PRT  
; ORGANISM: Lactobacillus rhamnosus  
US-09-971-536-52

Alignment Scores:  
Pred. No.: 2.28 Length: 495  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.38% Indels: 0  
DB: 9 Gaps: 0

US-09-714-865-15 (1-2172) x US-09-971-536-52 (1-495)

QY 1870 TATGTCATGCAATTGGCGTACTGTCGT 1899  
DB 316 TyrValHisArgIleGlyArgThrGlyArg 325

## RESULT 15

US-09-815-242-13491  
; Sequence 13491 Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13491  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13491

## Alignment Scores:

Pred. No.:	2.26	Length:	524
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.38%	Indels:	0
DB:	10	Gaps:	0

US-09-714-865-15 (1-2172) x US-09-815-242-13491 (1-524)

QY 1870 TATGTCATGCAATTGGCGTACTGTCGT 1899  
DB 323 TyrValHisArgIleGlyArgThrGlyArg 332

Search completed: June 10, 2003, 17:22:48  
Job time : 53.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 17:05:01 ; Search time 18.5 Seconds

(without alignments)

6908.820 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 724  
Sequence: 1 atgggggagatgaagatcgga.....tagatgatgagcatggat 2172

## Scoring table:

OLIGO	Xgapex 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

## Command line parameters:

-MODEL=frame+ n2p.model -DEV=xld  
-O=/cgm2\_1/USPTO.spool/US09714865/runat.05062003.111834.26626/app.query.fasta.1.2311  
-DB=Issued\_PatentAA -OPMT=faetan -SUFFIX=oligo.ra1 -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=1000 -DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USR=US09714865 @CGN 1.1 16 @runat.05062003.111834.26626 -NCPU=6 -ICPU=3  
-NO\_MMAP -IARGEDUBRY -NEG\_SCORES=0 -WAIT -DSBBLCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6  
-Fgapext=7 -Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

## Database :

1: /cgm2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgm2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgm2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgm2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgm2\_6/ptodata/2/iaa/PTCUS\_COMB.pep:\*  
6: /cgm2\_6/ptodata/2/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	3.0	662	3	US-09-058-489-15
2	22	3.0	662	3	US-09-058-489-15
3	18	2.5	660	3	US-09-058-489-18
4	18	2.5	660	3	US-09-058-489-18
5	11	1.5	648	4	US-09-183-706-43
6	11	1.5	648	4	US-09-183-706-43
7	10	1.4	149	4	US-08-679-493A-85
8	9	1.3	191	4	US-09-393-627B-25
9	9	1.3	543	2	US-08-224-482-4
10	9	1.2	746	4	US-09-149-934-4
11	9	1.3	870	2	US-09-010-928B-2
12	9	1.2	1261	4	US-09-208-742-4

13	9	1.2	1261	4	US-09-332-295-2	Sequence 2, Appl1
14	9	1.2	1261	4	US-09-709-979-2	Sequence 2, Appl1
15	8	1.1	17	1	US-08-264-093-24	Sequence 24, Appl1
16	8	1.1	78	2	US-08-245-511-20	Sequence 20, Appl1
17	8	1.1	78	2	US-08-600-993A-20	Sequence 20, Appl1
18	8	1.1	86	4	US-09-009-816-8	Sequence 8, Appl1
19	8	1.1	112	2	US-08-678-194-6	Sequence 6, Appl1
20	8	1.1	112	4	US-08-990-011-6	Sequence 6, Appl1
21	8	1.1	112	4	US-09-262-724-6	Sequence 6, Appl1
22	8	1.1	113	1	US-08-264-093-10	Sequence 10, Appl1
23	8	1.1	166	4	US-09-134-001C-4279	Sequence 4279, Ap
24	8	1.1	191	4	US-09-134-001C-4279	Sequence 22, Appl1
25	8	1.1	271	2	US-08-467-265-2	Sequence 2, Appl1
26	8	1.1	271	4	US-08-467-265-2	Sequence 2, Appl1
27	8	1.1	271	4	US-09-407-891-2	Sequence 2, Appl1
28	8	1.1	367	4	US-09-009-816-2	Sequence 2, Appl1
29	8	1.1	479	4	US-09-149-934-3	Sequence 3, Appl1
30	8	1.1	528	4	US-09-134-001C-4262	Sequence 4262, Ap
31	8	1.1	535	4	US-08-983-035A-38	Sequence 38, Appl1
32	8	1.1	547	3	US-09-176-657-1	Sequence 1, Appl1
33	8	1.1	752	2	US-08-896-590A-2	Sequence 2, Appl1
34	8	1.1	752	2	US-08-896-590A-4	Sequence 2, Appl1
35	8	1.1	752	4	US-09-185-832-2	Sequence 2, Appl1
36	8	1.1	752	4	US-09-185-832-4	Sequence 4, Appl1
37	8	1.1	859	4	US-09-149-934-1	Sequence 1, Appl1
38	7	1.0	19	1	US-07-908-317-7	Sequence 7, Appl1
39	7	1.0	19	4	US-08-630-915A-170	Sequence 170, App
40	7	1.0	19	5	PCT-US93-06171-7	Sequence 7, Appl1
41	7	1.0	22	4	US-09-177-249-302	Sequence 302, App
42	7	1.0	23	4	US-09-098-901-7	Sequence 7, Appl1
43	7	1.0	23	4	US-09-098-901-11	Sequence 11, Appl1
44	7	1.0	41	2	US-07-921-447-10	Sequence 10, Appl1
45	7	1.0	48	2	US-08-370-156-9	Sequence 9, Appl1
46	7	1.0	48	3	US-08-990-065-22	Sequence 22, Appl1
47	7	1.0	48	4	US-08-975-084-4	Sequence 4, Appl1
48	7	1.0	53	4	US-09-288-143-81	Sequence 8, Appl1
49	7	1.0	56	6	5496550-10	Patent No. 5496550
50	7	1.0	61	4	US-09-227-357-621	Sequence 621, App
51	7	1.0	64	4	US-09-461-697-267	Sequence 267, App
52	7	1.0	68	2	US-08-484-434-21	Sequence 21, Appl1
53	7	1.0	75	4	US-09-461-697-265	Sequence 265, App
54	7	1.0	76	4	US-09-605-785-575	Sequence 575, App
55	7	1.0	78	2	US-08-245-511-39	Sequence 39, Appl1
56	7	1.0	78	2	US-08-600-993A-39	Sequence 39, Appl1
57	7	1.0	85	4	US-09-430-564-8	Sequence 8, Appl1
58	7	1.0	85	4	US-08-245-511-38	Sequence 38, Appl1
59	7	1.0	86	2	US-08-600-993A-38	Sequence 38, Appl1
60	7	1.0	88	4	US-08-936-165A-513	Sequence 513, App
61	7	1.0	91	1	US-08-469-202-18	Sequence 18, Appl1
62	7	1.0	109	4	US-09-134-001C-3265	Sequence 3265, Ap
63	7	1.0	138	4	US-09-134-001C-2881	Sequence 2881, Ap
64	7	1.0	138	6	5268359-2	Patent No. 5268359
65	7	1.0	136	2	US-08-665-647-15	Sequence 15, Appl1
66	7	1.0	137	1	US-08-392-828C-38	Sequence 38, Appl1
67	7	1.0	137	3	US-09-330-945-38	Sequence 38, Appl1
68	7	1.0	144	1	US-08-133-979A-21	Sequence 21, Appl1
69	7	1.0	144	2	US-08-436-890-21	Sequence 21, Appl1
70	7	1.0	144	2	US-08-431-213-21	Sequence 21, Appl1
71	7	1.0	144	2	US-08-431-213-21	Sequence 21, Appl1
72	7	1.0	155	2	US-08-530-010-14	Sequence 14, Appl1
73	7	1.0	155	2	US-08-484-101B-14	Sequence 14, Appl1
74	7	1.0	155	4	US-08-714-524D-14	Sequence 14, Appl1
75	7	1.0	170	4	US-08-259-451-19	Sequence 19, Appl1
76	7	1.0	171	5	PCT-US95-04910-12	Sequence 12, Appl1
77	7	1.0	176	4	US-09-411-722-1	Sequence 1, Appl1
78	7	1.0	177	4	US-09-318-443-4	Sequence 4, Appl1
79	7	1.0	182	1	US-09-724-864-48	Sequence 48, Appl1
80	7	1.0	182	1	US-08-086-428B-62	Sequence 62, Appl1
81	7	1.0	192	1	US-08-086-428B-64	Sequence 64, Appl1
82	7	1.0	192	2	US-08-468-570-62	Sequence 62, Appl1
83	7	1.0	192	2	US-08-468-570-64	Sequence 64, Appl1
84	7	1.0	192	2	US-08-290-665A-62	Sequence 62, Appl1
85	7	1.0	192	2	US-08-290-665A-64	Sequence 64, Appl1

C 86	7	1.0	192	5	PCT-US95-10398-62	Sequence 62, Appl	C 159	7	1.0	378	2	US-09-055-097-1	Sequence 1, Appl
C 87	7	1.0	192	5	PCT-US95-10398-64	Sequence 64, Appl	C 160	7	1.0	378	4	US-09-186-489-2	Sequence 2, Appl
C 88	7	1.0	199	3	US-08-506-553C-8	Sequence 8, Appl	C 161	7	1.0	387	4	US-09-175-928-2	Sequence 2, Appl
C 89	7	1.0	205	4	US-09-637-481-2	Sequence 2, Appl	C 162	7	1.0	391	4	US-09-594-669-4	Sequence 4, Appl
C 90	7	1.0	205	4	US-09-723-428-2	Sequence 2, Appl	C 163	7	1.0	393	2	US-09-026-587-3	Sequence 3, Appl
C 91	7	1.0	208	2	US-08-606-143-41	Sequence 41, Appl	C 164	7	1.0	393	2	US-09-227-420-3	Sequence 3, Appl
C 92	7	1.0	210	4	US-08-612-973-14	Sequence 14, Appl	C 165	7	1.0	397	1	US-08-647-928-8	Sequence 8, Appl
C 93	7	1.0	210	4	US-08-927-597-14	Sequence 14, Appl	C 166	7	1.0	401	2	US-08-549-004-5	Sequence 5, Appl
C 94	7	1.0	216	3	US-08-506-553C-23	Sequence 23, Appl	C 167	7	1.0	401	3	US-08-517-802-3	Sequence 3, Appl
C 95	7	1.0	224	2	US-08-616-857-2	Sequence 2, Appl	C 168	7	1.0	401	4	US-09-051-802A-5	Sequence 5, Appl
C 96	7	1.0	228	4	US-09-724-864-42	Sequence 42, Appl	C 169	7	1.0	403	2	US-08-533-669-10	Sequence 10, Appl
C 97	7	1.0	244	1	US-08-288-699A-3	Sequence 3, Appl	C 170	7	1.0	403	2	US-08-607-509-2	Sequence 2, Appl
C 98	7	1.0	244	2	US-08-878-283-3	Sequence 3, Appl	C 171	7	1.0	403	2	US-08-607-509-4	Sequence 4, Appl
C 99	7	1.0	244	4	US-09-182-616-3	Sequence 3, Appl	C 172	7	1.0	403	2	US-08-454-303-2	Sequence 2, Appl
C 100	7	1.0	245	4	US-09-102-548B-9	Sequence 9, Appl	C 173	7	1.0	403	2	US-08-634-642-2	Sequence 2, Appl
C 101	7	1.0	245	4	US-09-105-058C-8	Sequence 8, Appl	C 174	7	1.0	403	2	US-08-634-642-4	Sequence 4, Appl
C 102	7	1.0	247	3	US-09-167-434-3	Sequence 3, Appl	C 175	7	1.0	403	3	US-08-989-370-2	Sequence 3, Appl
C 103	7	1.0	247	3	US-09-167-434-4	Sequence 4, Appl	C 176	7	1.0	403	3	US-08-989-370-4	Sequence 4, Appl
C 104	7	1.0	247	3	US-08-853-755-3	Sequence 3, Appl	C 177	7	1.0	403	4	US-09-183-861-10	Sequence 10, Appl
C 105	7	1.0	247	3	US-08-853-755-4	Sequence 4, Appl	C 178	7	1.0	403	4	US-09-022-765-10	Sequence 10, Appl
C 106	7	1.0	247	5	PCT-US94-10257A-2	Sequence 2, Appl	C 179	7	1.0	403	5	PCT-US95-0506A-2	Sequence 2, Appl
C 107	7	1.0	257	1	US-08-451-947-10	Sequence 10, Appl	C 180	7	1.0	404	1	US-07-602-824-4	Sequence 4, Appl
C 108	7	1.0	257	2	US-08-424-826A-10	Sequence 10, Appl	C 181	7	1.0	404	1	US-07-602-808-4	Sequence 4, Appl
C 109	7	1.0	257	3	US-08-928-694-10	Sequence 10, Appl	C 182	7	1.0	404	1	US-07-963-851-4	Sequence 4, Appl
C 110	7	1.0	257	5	PCT-US91-06950-10	Sequence 10, Appl	C 183	7	1.0	404	1	US-08-261-578-1	Sequence 1, Appl
C 111	7	1.0	259	3	US-09-082-090-2	Sequence 2, Appl	C 184	7	1.0	405	4	US-09-594-669-12	Sequence 12, Appl
C 112	7	1.0	268	1	US-08-431-387-4	Sequence 2, Appl	C 185	7	1.0	405	4	US-09-594-669-12	Sequence 12, Appl
C 113	7	1.0	274	4	US-09-723-430-2	Sequence 2, Appl	C 186	7	1.0	408	1	US-07-951-715A-21	Sequence 21, Appl
C 114	7	1.0	274	4	US-09-632-155-2	Sequence 2, Appl	C 187	7	1.0	408	2	US-08-459-48A-21	Sequence 21, Appl
C 115	7	1.0	274	4	US-09-724-518-2	Sequence 2, Appl	C 188	7	1.0	408	3	US-08-459-48A-21	Sequence 21, Appl
C 116	7	1.0	276	3	US-08-506-553C-26	Sequence 26, Appl	C 189	7	1.0	408	3	US-08-459-48A-21	Sequence 21, Appl
C 117	7	1.0	281	1	US-08-040-548-3	Sequence 3, Appl	C 190	7	1.0	408	3	US-08-459-404B-21	Sequence 21, Appl
C 118	7	1.0	281	1	US-08-466-344-3	Sequence 3, Appl	C 191	7	1.0	408	4	US-09-547-422-21	Sequence 21, Appl
C 119	7	1.0	300	3	US-08-765-856-2	Sequence 2, Appl	C 192	7	1.0	409	4	US-09-572-191-6	Sequence 6, Appl
C 120	7	1.0	300	4	US-08-935-009A-2	Sequence 2, Appl	C 193	7	1.0	409	4	US-09-723-262-6	Sequence 6, Appl
C 121	7	1.0	300	4	US-09-105-058C-4	Sequence 4, Appl	C 194	7	1.0	409	4	US-09-723-219-6	Sequence 6, Appl
C 122	7	1.0	300	4	US-09-105-058C-6	Sequence 6, Appl	C 195	7	1.0	411	4	US-08-460-900C-10	Sequence 10, Appl
C 123	7	1.0	301	3	US-08-829-525-24	Sequence 24, Appl	C 196	7	1.0	411	4	US-09-318-443-6	Sequence 6, Appl
C 124	7	1.0	301	4	US-08-609-583A-24	Sequence 24, Appl	C 197	7	1.0	411	4	US-09-318-443-8	Sequence 8, Appl
C 125	7	1.0	301	4	US-08-937-339-24	Sequence 24, Appl	C 198	7	1.0	411	4	US-08-954-698-10	Sequence 10, Appl
C 126	7	1.0	301	4	US-09-310-367-24	Sequence 24, Appl	C 199	7	1.0	411	4	US-08-957-874-10	Sequence 10, Appl
C 127	7	1.0	301	4	US-09-032-337-24	Sequence 24, Appl	C 200	7	1.0	411	4	US-09-335-256-19	Sequence 19, Appl
C 128	7	1.0	302	4	US-08-765-856-4	Sequence 4, Appl	C 201	7	1.0	414	5	PCT-US92-06840-2	Sequence 2, Appl
C 129	7	1.0	302	4	US-08-935-009A-4	Sequence 4, Appl	C 202	7	1.0	415	4	US-09-318-443-2	Sequence 2, Appl
C 130	7	1.0	305	1	US-09-134-001C-5632	Sequence 5632, Ap	C 203	7	1.0	418	2	US-09-026-587-1	Sequence 1, Appl
C 131	7	1.0	306	1	US-08-279-754-2	Sequence 2, Appl	C 204	7	1.0	418	2	US-09-227-420-1	Sequence 1, Appl
C 132	7	1.0	306	5	PCT-US95-09052-2	Sequence 2, Appl	C 205	7	1.0	430	2	US-08-924-440-2	Sequence 2, Appl
C 133	7	1.0	317	1	US-08-221-750A-13	Sequence 13, Appl	C 206	7	1.0	430	4	US-09-594-669-2	Sequence 2, Appl
C 134	7	1.0	326	2	US-09-055-097-6	Sequence 6, Appl	C 207	7	1.0	434	4	US-09-594-669-10	Sequence 10, Appl
C 135	7	1.0	335	4	US-08-736-915-2	Sequence 4, Appl	C 208	7	1.0	448	2	US-08-959-749-2	Sequence 2, Appl
C 136	7	1.0	337	4	US-09-641-806-4	Sequence 4, Appl	C 209	7	1.0	448	4	US-09-351-497-2	Sequence 2, Appl
C 137	7	1.0	341	4	US-09-724-517-4	Sequence 4, Appl	C 210	7	1.0	449	4	US-09-293-505-29	Sequence 29, Appl
C 138	7	1.0	341	4	US-09-641-807A-4	Sequence 4, Appl	C 211	7	1.0	455	4	US-09-134-001C-3653	Sequence 3653, Ap
C 139	7	1.0	341	4	US-09-723-096-4	Sequence 4, Appl	C 212	7	1.0	464	1	US-07-951-715A-22	Sequence 22, Appl
C 140	7	1.0	342	4	US-09-641-806-2	Sequence 2, Appl	C 213	7	1.0	464	2	US-08-459-448A-22	Sequence 22, Appl
C 141	7	1.0	347	3	US-08-445-515-58	Sequence 58, Appl	C 214	7	1.0	464	3	US-08-459-448A-22	Sequence 22, Appl
C 142	7	1.0	348	3	US-08-445-515-56	Sequence 56, Appl	C 215	7	1.0	464	3	US-08-459-444A-22	Sequence 22, Appl
C 143	7	1.0	349	3	US-09-009-620-2	Sequence 2, Appl	C 216	7	1.0	464	3	US-08-459-444-2	Sequence 3, Appl
C 144	7	1.0	350	1	US-08-458-067-2	Sequence 2, Appl	C 217	7	1.0	464	4	US-09-547-422-22	Sequence 22, Appl
C 145	7	1.0	350	5	PCT-US96-07795-2	Sequence 2, Appl	C 218	7	1.0	467	4	US-09-126-420A-17	Sequence 17, Appl
C 146	7	1.0	350	5	PCT-US96-07796-2	Sequence 2, Appl	C 219	7	1.0	469	2	US-08-477-451-23	Sequence 23, Appl
C 147	7	1.0	357	4	US-09-718-841-4	Sequence 4, Appl	C 220	7	1.0	473	4	US-09-594-669-8	Sequence 8, Appl
C 148	7	1.0	357	4	US-09-718-810-4	Sequence 4, Appl	C 221	7	1.0	473	4	US-09-594-669-8	Sequence 8, Appl
C 149	7	1.0	359	4	US-09-444-336-1	Sequence 1, Appl	C 222	7	1.0	480	2	US-08-828-488-8	Sequence 8, Appl
C 150	7	1.0	361	3	US-09-034-985-2	Sequence 2, Appl	C 223	7	1.0	480	3	US-09-189-035-5	Sequence 5, Appl
C 151	7	1.0	361	4	US-09-596-541-2	Sequence 2, Appl	C 224	7	1.0	480	4	US-09-382-086-5	Sequence 5, Appl
C 152	7	1.0	362	4	US-09-594-669-6	Sequence 6, Appl	C 225	7	1.0	480	4	US-09-299-689A-8	Sequence 8, Appl
C 153	7	1.0	362	4	US-09-444-336-2	Sequence 2, Appl	C 226	7	1.0	482	4	US-09-430-564A-2	Sequence 4, Appl
C 154	7	1.0	369	4	US-09-596-541-6	Sequence 6, Appl	C 227	7	1.0	483	3	US-09-027-166-7	Sequence 7, Appl
C 155	7	1.0	370	4	US-09-596-541-4	Sequence 4, Appl	C 228	7	1.0	485	1	US-08-453-956-15	Sequence 15, Appl
C 156	7	1.0	375	4	US-09-572-191-4	Sequence 4, Appl	C 229	7	1.0	485	1	US-08-086-631-15	Sequence 15, Appl
C 157	7	1.0	375	4	US-09-723-262-4	Sequence 4, Appl	C 230	7	1.0	485	2	US-08-452-930-15	Sequence 15, Appl
C 158	7	1.0	375	4	US-09-723-219-4	Sequence 4, Appl	C 231	7	1.0	485	5	PCT-US93-08174-15	Sequence 15, Appl

C 232	7	1.0	488	2	US-08-928-692-10	Sequence 10, Appl	C 305	7	1.0	758	1	US-08-289-112-2	Sequence 2, Appl
C 233	7	1.0	488	4	US-09-339-972-10	Sequence 10, Appl	C 306	7	1.0	763	2	US-08-677-862-2	Sequence 2, Appl
C 234	7	1.0	492	1	US-07-794-393-4	Sequence 4, Appl	C 307	7	1.0	763	2	US-09-252-571-1	Sequence 2, Appl
C 235	7	1.0	492	1	US-08-001-711-4	Sequence 4, Appl	C 308	7	1.0	763	3	US-09-434-065-2	Sequence 2, Appl
C 236	7	1.0	492	4	US-09-152-060-72	Sequence 72, Appl	C 309	7	1.0	763	4	US-08-789-275-5	Sequence 4, Appl
C 237	7	1.0	495	2	US-08-749-992-3	Sequence 3, Appl	C 310	7	1.0	763	4	US-08-789-275-5	Sequence 5, Appl
C 238	7	1.0	495	2	US-09-430-564-3	Sequence 3, Appl	C 311	7	1.0	812	4	US-09-166-550-12	Sequence 1, Appl
C 239	7	1.0	496	2	US-08-224-482-2	Sequence 2, Appl	C 312	7	1.0	814	1	US-08-233-788A-42	Sequence 1, Appl
C 240	7	1.0	504	4	US-08-868-373-6	Sequence 6, Appl	C 313	7	1.0	822	2	US-08-939-002A-16	Sequence 1, Appl
C 241	7	1.0	517	3	US-09-189-760-2	Sequence 2, Appl	C 314	7	1.0	837	4	US-09-390-334-14	Sequence 14, Appl
C 242	7	1.0	517	3	US-09-514-422-2	Sequence 2, Appl	C 315	7	1.0	863	4	US-09-619-353-14	Sequence 14, Appl
C 243	7	1.0	518	3	US-09-043-123-2	Sequence 2, Appl	C 316	7	1.0	871	4	US-09-105-058C-20	Sequence 20, Appl
C 244	7	1.0	522	4	US-09-592-054-4	Sequence 4, Appl	C 317	7	1.0	872	4	US-09-177-650-2	Sequence 2, Appl
C 245	7	1.0	533	1	US-08-040-548-1	Sequence 1, Appl	C 318	7	1.0	911	4	US-08-461-562B-2	Sequence 2, Appl
C 246	7	1.0	533	1	US-08-466-344-1	Sequence 1, Appl	C 319	7	1.0	930	4	US-09-177-650-96	Sequence 96, Appl
C 247	7	1.0	533	6	5206152-2	Patent No. 5206152	C 320	7	1.0	934	1	US-08-464-486-7	Sequence 7, Appl
C 248	7	1.0	534	4	US-08-356-786-10	Sequence 10, Appl	C 321	7	1.0	934	1	US-08-463-308-7	Sequence 7, Appl
C 249	7	1.0	534	4	US-09-103-664A-2	Sequence 2, Appl	C 322	7	1.0	955	1	US-08-006-676B-1	Sequence 1, Appl
C 250	7	1.0	544	4	US-08-687-590-10	Sequence 30, Appl	C 323	7	1.0	955	1	US-08-282-845-2	Sequence 2, Appl
C 251	7	1.0	552	3	US-09-295-186-10	Sequence 10, Appl	C 324	7	1.0	955	2	US-08-428-414A-3	Sequence 3, Appl
C 252	7	1.0	553	2	US-08-943-087-2	Sequence 2, Appl	C 325	7	1.0	955	5	PCT-US94-00324-1	Sequence 1, Appl
C 253	7	1.0	553	2	US-08-943-087-14	Sequence 14, Appl	C 326	7	1.0	982	4	US-09-556-877-176	Sequence 176, App
C 254	7	1.0	553	2	US-08-943-087-16	Sequence 16, Appl	C 327	7	1.0	982	4	US-09-620-412C-176	Sequence 176, App
C 255	7	1.0	553	2	US-08-943-087-18	Sequence 18, Appl	C 328	7	1.0	1006	4	US-09-556-877-190	Sequence 190, App
C 256	7	1.0	553	2	US-08-943-087-20	Sequence 20, Appl	C 329	7	1.0	1006	4	US-09-620-412C-190	Sequence 190, App
C 257	7	1.0	553	2	US-08-943-087-22	Sequence 22, Appl	C 330	7	1.0	1093	3	US-08-545-860D-55	Sequence 55, Appl
C 258	7	1.0	553	2	US-08-943-087-24	Sequence 24, Appl	C 331	7	1.0	1093	5	PCT-US94-00496-55	Sequence 55, Appl
C 259	7	1.0	553	2	US-08-943-087-26	Sequence 26, Appl	C 332	7	1.0	1164	4	US-09-457-708-2	Sequence 2, Appl
C 260	7	1.0	553	2	US-08-943-087-28	Sequence 28, Appl	C 333	7	1.0	1231	4	US-09-071-035-420	Sequence 420, App
C 261	7	1.0	553	2	US-08-943-087-30	Sequence 30, Appl	C 334	7	1.0	1232	4	US-09-582-054-2	Sequence 2, Appl
C 262	7	1.0	553	2	US-08-943-087-32	Sequence 32, Appl	C 335	7	1.0	1251	5	US-09-592-054-8	Sequence 8, Appl
C 263	7	1.0	553	2	US-08-943-087-34	Sequence 34, Appl	C 336	7	1.0	1251	5	PCT-US95-02251-3	Sequence 3, Appl
C 264	7	1.0	553	2	US-08-943-087-36	Sequence 36, Appl	C 337	7	1.0	1252	1	US-08-199-780-3	Sequence 3, Appl
C 265	7	1.0	553	2	US-08-943-087-38	Sequence 38, Appl	C 338	7	1.0	1252	2	US-08-316-650-3	Sequence 3, Appl
C 266	7	1.0	553	2	US-08-943-087-40	Sequence 40, Appl	C 339	7	1.0	1253	3	US-08-479-722B-4	Sequence 4, Appl
C 267	7	1.0	553	2	US-08-943-087-42	Sequence 42, Appl	C 340	7	1.0	1265	4	US-09-071-035-418	Sequence 4, Appl
C 268	7	1.0	553	2	US-08-943-087-44	Sequence 44, Appl	C 341	7	1.0	1279	4	US-09-724-517-2	Sequence 2, Appl
C 269	7	1.0	553	2	US-08-943-087-46	Sequence 46, Appl	C 342	7	1.0	1279	4	US-09-641-807A-2	Sequence 2, Appl
C 270	7	1.0	553	2	US-08-943-087-48	Sequence 48, Appl	C 343	7	1.0	1279	4	US-09-723-096-2	Sequence 2, Appl
C 271	7	1.0	553	3	US-09-083-351-2	Sequence 2, Appl	C 344	7	1.0	1306	4	US-08-989-299-7	Sequence 7, Appl
C 272	7	1.0	553	3	US-09-083-352-2	Sequence 2, Appl	C 345	7	1.0	1311	2	US-08-540-406-4	Sequence 4, Appl
C 273	7	1.0	560	3	US-09-176-657-7	Sequence 7, Appl	C 346	7	1.0	1311	3	US-08-656-055-4	Sequence 4, Appl
C 274	7	1.0	563	4	US-09-718-841-2	Sequence 2, Appl	C 347	7	1.0	1311	4	US-08-954-668-4	Sequence 4, Appl
C 275	7	1.0	563	4	US-09-718-841-2	Sequence 2, Appl	C 348	7	1.0	1311	4	US-08-918-658-4	Sequence 4, Appl
C 276	7	1.0	564	4	US-09-134-001C-3930	Sequence 3930, Ap	C 349	7	1.0	1311	5	PCT-US95-11323-4	Sequence 4, Appl
C 277	7	1.0	567	3	US-08-646-273-19	Sequence 19, Appl	C 350	7	1.0	1312	4	US-09-041-886-19	Sequence 19, Appl
C 278	7	1.0	567	1	US-09-134-001C-5646	Sequence 5646, Ap	C 351	7	1.0	1375	4	US-09-210-361-4	Sequence 4, Appl
C 279	7	1.0	572	1	US-08-160-861-3	Sequence 3, Appl	C 352	7	1.0	1375	4	US-09-740-274-4	Sequence 4, Appl
C 280	7	1.0	573	3	US-09-295-186-11	Sequence 11, Appl	C 353	7	1.0	1388	4	US-09-572-191-2	Sequence 2, Appl
C 281	7	1.0	582	1	US-08-261-086-2	Sequence 2, Appl	C 354	7	1.0	1388	4	US-09-723-262-2	Sequence 2, Appl
C 282	7	1.0	617	2	US-08-370-156-6	Sequence 6, Appl	C 355	7	1.0	1388	4	US-09-723-219-2	Sequence 2, Appl
C 283	7	1.0	617	3	US-08-814-095-6	Sequence 6, Appl	C 356	7	1.0	1620	1	US-08-542-163-2	Sequence 2, Appl
C 284	7	1.0	635	1	US-07-832-855-2	Sequence 2, Appl	C 357	7	1.0	1620	4	US-09-100-089-2	Sequence 2, Appl
C 285	7	1.0	635	4	US-08-176-320-2	Sequence 2, Appl	C 358	7	1.0	1620	4	US-09-670-827-2	Sequence 2, Appl
C 286	7	1.0	635	4	US-08-751-305-2	Sequence 2, Appl	C 359	7	1.0	1646	4	US-09-535-008-67	Sequence 67, Appl
C 287	7	1.0	655	1	US-08-148-910-12	Sequence 12, Appl	C 360	7	1.0	1647	4	US-09-535-008-75	Sequence 2, Appl
C 288	7	1.0	655	1	US-08-448-937A-12	Sequence 12, Appl	C 361	7	1.0	1649	4	US-09-535-008-75	Sequence 75, Appl
C 289	7	1.0	686	2	US-08-993-238-12	Sequence 12, Appl	C 362	7	1.0	1650	4	US-09-535-008-71	Sequence 71, Appl
C 290	7	1.0	695	1	US-08-164-839-8	Sequence 8, Appl	C 363	7	1.0	1663	2	US-08-793-126-1	Sequence 1, Appl
C 291	7	1.0	695	1	US-08-583-799-8	Sequence 8, Appl	C 364	7	1.0	1663	4	US-09-132-271-1	Sequence 1, Appl
C 292	7	1.0	696	1	US-08-164-839-10	Sequence 10, Appl	C 365	7	1.0	1663	4	US-09-142-334-22	Sequence 22, Appl
C 293	7	1.0	696	1	US-08-583-799-10	Sequence 10, Appl	C 366	7	1.0	1678	4	US-09-535-008-69	Sequence 69, Appl
C 294	7	1.0	703	3	US-08-646-273-25	Sequence 25, Appl	C 367	7	1.0	1679	4	US-09-535-008-65	Sequence 65, Appl
C 295	7	1.0	703	3	US-09-134-001C-3015	Sequence 3015, Ap	C 368	7	1.0	1681	4	US-09-535-008-77	Sequence 77, Appl
C 296	7	1.0	708	3	US-08-646-273-23	Sequence 23, Appl	C 369	7	1.0	1682	4	US-09-535-008-73	Sequence 73, Appl
C 297	7	1.0	722	4	US-09-105-058C-23	Sequence 23, Appl	C 370	7	1.0	1704	4	US-08-485-355B-40	Sequence 40, Appl
C 298	7	1.0	723	4	US-09-594-669-14	Sequence 14, Appl	C 371	7	1.0	1872	6	5386025-6	Patent No. 5386025
C 299	7	1.0	725	4	US-09-594-669-16	Sequence 16, Appl	C 372	7	1.0	1873	1	US-08-435-675B-4	Sequence 4, Appl
C 300	7	1.0	730	4	US-09-291-170A-3	Sequence 3, Appl	C 373	7	1.0	1873	4	US-08-336-257A-7	Sequence 7, Appl
C 301	7	1.0	730	4	US-09-724-884-3	Sequence 3, Appl	C 374	7	1.0	2318	1	US-09-091-219-24	Sequence 24, Appl
C 302	7	1.0	753	3	US-08-646-273-36	Sequence 36, Appl	C 375	7	1.0	2368	1	US-08-198-446B-15	Sequence 15, Appl
C 303	7	1.0	754	3	US-08-646-273-30	Sequence 30, Appl	C 376	7	1.0	2368	2	US-08-870-693-15	Sequence 15, Appl
C 304	7	1.0	757	4	US-09-177-650-69	Sequence 69, Appl	C 377	7	1.0	3177	2	US-08-477-451-4	Sequence 4, Appl

C 378	6	0.9	8	1	US-08-479-233-10	Sequence 10, Appl	C 451	6	0.9	14	4	US-09-248-061B-17	Sequence 17, Appl
C 379	6	0.9	8	5	PCT-US93-00643-10	Sequence 10, Appl	452	6	0.8	14	4	US-09-316-919-62	Sequence 62, Appl
C 380	6	0.9	9	2	US-08-448-108-88	Sequence 88, Appl	453	6	0.8	14	5	PCT-US93-12169-3	Sequence 3, Appl
C 381	6	0.9	9	2	US-08-801-319-2	Sequence 2, Appl	454	6	0.8	15	1	US-07-763-512-2	Sequence 2, Appl
C 382	6	0.8	10	3	US-08-455-073A-11	Sequence 11, Appl	455	6	0.8	15	1	US-08-137-800-5	Sequence 5, Appl
C 383	6	0.9	10	3	US-08-390-874C-2	Sequence 2, Appl	456	6	0.9	15	1	US-08-230-047-29	Sequence 29, Appl
C 384	6	0.8	10	3	US-08-899-595-8	Sequence 8, Appl	457	6	0.8	15	1	US-08-477-383-5	Sequence 5, Appl
C 385	6	0.8	10	3	US-09-199-242E-11	Sequence 11, Appl	458	6	0.8	15	1	US-08-487-174-5	Sequence 5, Appl
C 386	6	0.9	10	4	US-09-265-772-2	Sequence 2, Appl	459	6	0.8	15	1	US-08-480-750-5	Sequence 5, Appl
C 387	6	0.9	11	1	US-08-323-474-5	Sequence 5, Appl	460	6	0.8	15	2	US-08-373-190-4	Sequence 4, Appl
C 388	6	0.9	11	1	US-08-066-335C-13	Sequence 13, Appl	461	6	0.9	15	2	US-08-664-725-18	Sequence 18, Appl
C 389	6	0.9	11	1	US-08-641-704-1	Sequence 1, Appl	462	6	0.8	15	2	US-08-438-190A-4	Sequence 4, Appl
C 390	6	0.9	11	1	US-08-236-918A-17	Sequence 17, Appl	463	6	0.8	15	3	US-08-350-215-4	Sequence 4, Appl
C 391	6	0.8	11	1	US-08-469-582-6	Sequence 6, Appl	464	6	0.8	15	3	US-08-654-223-13	Sequence 13, Appl
C 392	6	0.8	11	1	US-08-469-582-7	Sequence 7, Appl	465	6	0.9	15	3	US-08-256-747C-26	Sequence 26, Appl
C 393	6	0.8	11	2	US-08-473-025-14	Sequence 14, Appl	466	6	0.8	15	3	US-08-630-172-38	Sequence 38, Appl
C 394	6	0.8	11	2	US-08-851-843A-186	Sequence 186, App	467	6	0.8	15	3	US-09-287-145A-4	Sequence 4, Appl
C 395	6	0.8	11	3	US-08-974-549A-305	Sequence 305, App	468	6	0.8	15	4	US-08-834-130A-26	Sequence 26, Appl
C 396	6	0.8	11	4	US-08-854-050-186	Sequence 186, App	469	6	0.8	15	4	US-08-602-999A-356	Sequence 356, App
C 397	6	0.8	11	4	US-09-430-323-186	Sequence 186, App	470	6	0.8	15	4	US-08-602-999A-356	Sequence 423, App
C 398	6	0.9	11	4	US-09-150-864A-17	Sequence 17, Appl	471	6	0.8	15	4	US-09-375-419-38	Sequence 38, Appl
C 399	6	0.9	11	5	PCT-US95-08533-1	Sequence 1, Appl	472	6	0.8	15	4	US-09-556-111-4	Sequence 4, Appl
C 400	6	0.8	12	1	US-08-469-582-9	Sequence 9, Appl	473	6	0.8	15	4	US-09-500-124-355	Sequence 355, App
C 401	6	0.8	12	2	US-08-950-866-8	Sequence 8, Appl	474	6	0.8	15	4	US-09-500-124-423	Sequence 423, App
C 402	6	0.8	12	2	US-08-473-025-13	Sequence 13, Appl	475	6	0.8	15	5	PCT-US93-06751-112	Sequence 112, App
C 403	6	0.8	12	2	US-08-224-591-7	Sequence 7, Appl	476	6	0.8	15	5	PCT-US94-01499-4	Sequence 4, Appl
C 404	6	0.8	12	2	US-08-392-338A-1	Sequence 1, Appl	477	6	0.8	15	5	PCT-US96-07962-2	Sequence 2, Appl
C 405	6	0.8	12	2	US-08-926-789-7	Sequence 7, Appl	478	6	0.8	16	1	US-07-994-469A-21	Sequence 21, Appl
C 406	6	0.8	12	2	US-08-818-253-41	Sequence 41, Appl	479	6	0.8	16	2	US-08-637-759B-465	Sequence 465, App
C 407	6	0.8	12	3	US-09-166-750-1	Sequence 1, Appl	480	6	0.9	16	2	US-08-859-201-18	Sequence 18, Appl
C 408	6	0.8	12	3	US-09-166-093-1	Sequence 1, Appl	481	6	0.8	16	2	US-08-859-201-20	Sequence 20, Appl
C 409	6	0.9	12	3	US-08-599-226-31	Sequence 31, Appl	482	6	0.8	16	2	US-08-597-803-3	Sequence 3, Appl
C 410	6	0.8	12	3	US-09-172-019-1	Sequence 1, Appl	483	6	0.9	16	2	US-08-672-345C-43	Sequence 43, Appl
C 411	6	0.8	12	3	US-09-166-093-1	Sequence 1, Appl	484	6	0.9	16	3	US-08-871-355A-465	Sequence 465, App
C 412	6	0.8	12	4	US-08-818-252-41	Sequence 41, Appl	485	6	0.9	16	3	US-08-788-231A-19	Sequence 19, Appl
C 413	6	0.9	12	4	US-09-125-098-31	Sequence 31, Appl	486	6	0.8	16	4	US-09-214-095D-43	Sequence 43, Appl
C 414	6	0.8	12	4	US-09-362-805-3	Sequence 3, Appl	487	6	0.9	16	4	US-09-514-302-5	Sequence 5, Appl
C 415	6	0.8	12	4	US-09-173-190-3	Sequence 3, Appl	488	6	0.9	16	4	US-09-514-302-5	Sequence 455, App
C 416	6	0.8	12	4	US-09-069-821-6	Sequence 6, Appl	489	6	0.9	16	6	US-09-201-945-465	Patent No. 5196511
C 417	6	0.8	12	5	US-09-316-919-57	Sequence 57, Appl	490	6	0.8	17	1	US-07-763-512-3	Sequence 3, Appl
C 418	6	0.8	12	5	PCT-US93-11138-7	Sequence 7, Appl	491	6	0.9	17	1	US-08-353-400-30	Sequence 30, Appl
C 419	6	0.8	13	1	US-07-763-512-5	Sequence 5, Appl	492	6	0.9	17	1	US-08-467-420A-10	Sequence 10, Appl
C 420	6	0.8	13	3	US-08-630-916A-7	Sequence 7, Appl	493	6	0.9	17	1	US-08-470-110A-10	Sequence 10, Appl
C 421	6	0.9	13	4	US-08-602-999A-78	Sequence 78, Appl	494	6	0.9	17	1	US-08-667-769A-10	Sequence 10, Appl
C 422	6	0.9	13	4	US-08-278-865-78	Sequence 78, Appl	495	6	0.9	17	2	US-08-940-371-10	Sequence 10, Appl
C 423	6	0.8	13	4	US-08-525-539A-66	Sequence 66, Appl	496	6	0.8	17	3	US-08-837-226-6	Sequence 6, Appl
C 424	6	0.9	13	4	US-09-500-124-78	Sequence 78, Appl	497	6	0.9	17	3	US-09-015-605-3	Sequence 3, Appl
C 425	6	0.8	13	4	US-08-469-260A-599	Sequence 599, App	498	6	0.9	17	3	US-08-617-647-10	Sequence 10, Appl
C 426	6	0.8	14	1	US-07-994-469A-20	Sequence 20, Appl	499	6	0.9	17	4	US-08-974-549A-61	Sequence 61, Appl
C 427	6	0.8	14	1	US-08-469-582-3	Sequence 3, Appl	500	6	0.9	17	4	US-07-987-264-4	Sequence 4, Appl
C 428	6	0.9	14	1	US-08-484-969-2	Sequence 2, Appl	501	6	0.9	17	4	US-09-265-653-19	Sequence 19, Appl
C 429	6	0.9	14	1	US-08-473-627-2	Sequence 2, Appl	502	6	0.9	17	4	US-09-265-653-22	Sequence 22, Appl
C 430	6	0.9	14	1	US-08-388-463-2	Sequence 2, Appl	503	6	0.9	17	5	PCT-US93-08435-35	Sequence 35, Appl
C 431	6	0.9	14	2	US-08-448-418-89	Sequence 89, Appl	504	6	0.9	17	5	PCT-US93-08435-35	Sequence 35, Appl
C 432	6	0.8	14	2	US-08-373-190-5	Sequence 3, Appl	505	6	0.9	17	5	PCT-US95-17082A-10	Sequence 10, Appl
C 433	6	0.8	14	2	US-08-473-025-8	Sequence 8, Appl	506	6	0.8	18	1	US-07-725-331-42	Sequence 42, Appl
C 434	6	0.8	14	2	US-08-438-190A-3	Sequence 3, Appl	507	6	0.8	18	1	US-07-725-331-42	Sequence 33, Appl
C 435	6	0.8	14	2	US-08-438-190A-5	Sequence 5, Appl	508	6	0.8	18	1	US-08-207-169A-7	Sequence 7, Appl
C 436	6	0.8	14	2	US-08-818-253-46	Sequence 46, Appl	509	6	0.9	18	1	US-08-894-339-7	Sequence 7, Appl
C 437	6	0.8	14	2	US-08-350-215-3	Sequence 3, Appl	510	6	0.8	18	2	US-08-466-860-20	Sequence 20, Appl
C 438	6	0.8	14	3	US-08-837-226-8	Sequence 8, Appl	511	6	0.8	18	2	US-08-472-040A-20	Sequence 20, Appl
C 439	6	0.8	14	3	US-08-837-226-8	Sequence 8, Appl	512	6	0.8	18	3	US-09-306-044-7	Sequence 7, Appl
C 440	6	0.9	14	3	US-09-287-145A-3	Sequence 3, Appl	513	6	0.8	18	4	US-08-471-209-20	Sequence 20, Appl
C 441	6	0.8	14	3	US-09-287-145A-3	Sequence 3, Appl	514	6	0.8	18	4	US-08-471-209-20	Sequence 20, Appl
C 442	6	0.8	14	3	US-09-287-145A-5	Sequence 5, Appl	515	6	0.8	18	4	US-08-471-209-20	Sequence 20, Appl
C 443	6	0.8	14	3	US-09-046-992-6	Sequence 6, Appl	516	6	0.8	18	5	PCT-US91-05047-33	Sequence 33, Appl
C 444	6	0.8	14	4	US-08-818-253-46	Sequence 46, Appl	517	6	0.8	18	5	PCT-US91-05047-42	Sequence 42, Appl
C 445	6	0.8	14	4	US-08-928-213B-167	Sequence 167, App	518	6	0.9	19	1	US-08-457-274A-3	Sequence 3, Appl
C 446	6	0.8	14	4	US-09-362-805-8	Sequence 8, Appl	519	6	0.9	19	2	US-08-652-450A-9	Sequence 9, Appl
C 447	6	0.8	14	4	US-09-173-190-8	Sequence 8, Appl	520	6	0.9	19	2	US-08-448-418-90	Sequence 90, Appl
C 448	6	0.8	14	4	US-09-556-111-3	Sequence 3, Appl	521	6	0.9	19	4	US-08-630-915A-169	Sequence 169, App
C 449	6	0.8	14	4	US-09-556-111-5	Sequence 5, Appl	522	6	0.9	19	4	US-09-225-322B-3	Sequence 3, Appl
C 450	6	0.9	14	4	US-09-400-653A-33	Sequence 33, Appl	523	6	0.9	19	5	PCT-US95-05758-3	Sequence 3, Appl



524	6	0.8	20	1	US-07-955-041-2	Sequence 2, Appl	597	6	0.8	23	5	PCT-US93-08528-93	Sequence 93, Appl
525	6	0.9	20	1	US-08-103-742-32	Sequence 32, Appl	598	6	0.9	24	3	US-08-592-500-25	Sequence 25, Appl
526	6	0.8	20	1	US-07-956-848A-44	Sequence 44, Appl	599	6	0.9	24	3	US-08-195-006-25	Sequence 25, Appl
527	6	0.8	20	1	US-08-227-455-2	Sequence 2, Appl	600	6	0.8	24	4	US-09-098-901-9	Sequence 9, Appl
528	6	0.9	20	1	US-07-678-974D-35	Sequence 35, Appl	601	6	0.9	24	5	PCT-US94-0764A-25	Sequence 25, Appl
529	6	0.8	20	1	US-08-472-482-2	Sequence 2, Appl	602	6	0.9	25	1	US-08-240-514-56	Sequence 56, Appl
530	6	0.8	20	1	US-08-467-069-2	Sequence 2, Appl	603	6	0.9	25	2	US-08-612-302A-56	Sequence 56, Appl
531	6	0.8	20	1	US-08-471-956-44	Sequence 44, Appl	604	6	0.9	25	2	US-08-902-516-4	Sequence 4, Appl
532	6	0.9	20	2	US-08-945-168-40	Sequence 40, Appl	605	6	0.9	25	2	US-08-902-516-10	Sequence 10, Appl
533	6	0.9	20	3	US-08-654-623-15	Sequence 15, Appl	606	6	0.9	25	4	US-08-278-865-45	Sequence 45, Appl
534	6	0.8	20	3	US-08-256-747C-12	Sequence 12, Appl	607	6	0.9	25	4	US-09-305-984A-29	Sequence 29, Appl
535	6	0.8	20	4	US-08-834-130A-12	Sequence 12, Appl	608	6	0.9	25	4	US-09-073-541A-29	Sequence 29, Appl
536	6	0.9	20	4	US-09-136-293-2	Sequence 2, Appl	609	6	0.9	25	5	PCT-US94-04361-36	Sequence 36, Appl
537	6	0.9	20	4	US-08-505-250-22	Sequence 22, Appl	610	6	0.9	26	1	US-07-942-245-328	Sequence 328, App
538	6	0.9	20	4	US-08-918-288-51	Sequence 51, Appl	611	6	0.8	26	1	US-07-942-245-425	Sequence 425, App
539	6	0.9	20	4	US-08-918-288-56	Sequence 56, Appl	612	6	0.9	26	1	US-07-942-245-487	Sequence 487, App
540	6	0.9	20	4	US-09-282-357-51	Sequence 51, Appl	613	6	0.9	26	1	US-07-942-245-490	Sequence 490, App
541	6	0.9	20	4	US-09-282-357-56	Sequence 56, Appl	614	6	0.9	26	3	US-08-630-172-40	Sequence 40, Appl
542	6	0.9	20	4	US-08-209-603B-2	Sequence 2, Appl	615	6	0.9	26	3	US-08-513-968-53	Sequence 53, Appl
543	6	0.9	20	4	US-08-505-250-22	Sequence 22, Appl	616	6	0.8	26	4	US-08-602-999A-45	Sequence 45, Appl
544	6	0.9	21	1	US-07-715-397A-2	Sequence 2, Appl	617	6	0.9	26	4	US-08-905-223-289	Sequence 289, App
545	6	0.9	21	1	US-07-715-397A-3	Sequence 3, Appl	618	6	0.9	26	4	US-09-375-419-40	Sequence 40, Appl
546	6	0.8	21	1	US-07-593-657-1	Sequence 1, Appl	619	6	0.9	26	4	US-09-336-536-59	Sequence 59, Appl
547	6	0.9	21	1	US-08-060-833-2	Sequence 2, Appl	620	6	0.9	26	4	US-09-500-124-45	Sequence 45, Appl
548	6	0.9	21	1	US-08-060-833-3	Sequence 3, Appl	621	6	0.8	27	1	US-07-927-071-6	Sequence 6, Appl
549	6	0.9	21	1	US-08-218-025A-124	Sequence 124, App	622	6	0.9	27	2	US-08-652-450A-7	Sequence 7, Appl
550	6	0.9	21	1	US-08-419-824-12	Sequence 12, Appl	623	6	0.9	27	2	US-08-652-450A-6	Sequence 6, Appl
551	6	0.9	21	1	US-08-419-824-2	Sequence 2, Appl	624	6	0.9	27	2	US-08-652-450A-7	Sequence 7, Appl
552	6	0.9	21	1	US-08-419-824-3	Sequence 3, Appl	625	6	0.8	27	2	US-08-833-546-8	Sequence 8, Appl
553	6	0.9	21	2	US-08-845-526-34	Sequence 34, Appl	626	6	0.9	27	3	US-08-654-623-23	Sequence 23, Appl
554	6	0.9	21	3	US-08-484-905-14	Sequence 14, Appl	627	6	0.8	27	3	US-08-256-747C-46	Sequence 46, Appl
555	6	0.9	21	3	US-08-481-985B-14	Sequence 14, Appl	628	6	0.8	27	4	US-08-834-130A-46	Sequence 46, Appl
556	6	0.9	21	3	US-08-848-580-2	Sequence 2, Appl	629	6	0.8	27	4	US-09-388-664-8	Sequence 8, Appl
557	6	0.9	21	3	US-08-848-580-3	Sequence 3, Appl	630	6	0.9	27	4	US-09-327-357-276	Sequence 276, App
558	6	0.9	21	3	US-08-848-580-6	Sequence 6, Appl	631	6	0.9	27	6	516318-17	Patent No. 516318
559	6	0.9	21	3	US-08-848-580-7	Sequence 7, Appl	632	6	0.9	28	1	US-08-164-151-22	Sequence 22, Appl
560	6	0.9	21	3	US-08-848-580-13	Sequence 13, Appl	633	6	0.8	28	1	US-08-459-568-58	Sequence 58, Appl
561	6	0.8	21	4	US-08-370-476-14	Sequence 14, Appl	634	6	0.8	28	2	US-08-399-411-58	Sequence 58, Appl
562	6	0.9	21	4	US-08-602-999A-427	Sequence 427, App	635	6	0.8	28	3	US-09-045-632-95	Sequence 95, Appl
563	6	0.8	21	4	US-09-500-124-427	Sequence 427, App	636	6	0.8	28	3	US-08-516-859A-58	Sequence 58, Appl
564	6	0.9	21	5	PCT-US92-04537-34	Sequence 34, Appl	637	6	0.8	28	4	US-08-957-001B-25	Sequence 25, Appl
565	6	0.9	21	5	PCT-US92-04537-2	Sequence 2, Appl	638	6	0.8	28	4	US-09-496-301-25	Sequence 25, Appl
566	6	0.9	21	5	PCT-US92-04537-3	Sequence 3, Appl	639	6	0.8	28	4	US-09-586-472-58	Sequence 58, Appl
567	6	0.9	21	5	PCT-US93-04537-8	Sequence 8, Appl	640	6	0.8	28	4	US-09-528-706-58	Sequence 58, Appl
568	6	0.9	21	5	PCT-US93-04537-9	Sequence 9, Appl	641	6	0.9	28	5	PCT-US95-04611-2	Sequence 2, Appl
569	6	0.8	21	5	5164481-4	Patent No. 5164481	642	6	0.8	29	1	US-07-729-099-4	Sequence 4, Appl
570	6	0.9	22	1	US-08-197-496A-6	Sequence 6, Appl	643	6	0.8	29	1	US-08-257-392-4	Sequence 4, Appl
571	6	0.9	22	1	US-08-275-370-6	Sequence 6, Appl	644	6	0.9	29	1	US-08-393-985-33	Sequence 33, Appl
572	6	0.9	22	1	US-08-367-968-6	Sequence 6, Appl	645	6	0.8	29	1	US-07-960-510-4	Sequence 4, Appl
573	6	0.9	22	2	US-08-665-484-6	Sequence 6, Appl	646	6	0.8	29	1	US-07-960-510-5	Sequence 5, Appl
574	6	0.9	22	2	US-08-792-553-16	Sequence 16, Appl	647	6	0.8	29	1	US-07-960-510-6	Sequence 6, Appl
575	6	0.9	22	4	US-09-430-564-9	Sequence 9, Appl	648	6	0.8	29	1	US-07-960-510-7	Sequence 7, Appl
576	6	0.9	22	4	US-09-430-564-15	Sequence 15, Appl	649	6	0.8	29	3	US-08-770-035-4	Sequence 4, Appl
577	6	0.8	23	1	US-08-118-270-93	Sequence 93, Appl	650	6	0.8	29	3	US-08-256-747C-28	Sequence 28, Appl
578	6	0.8	23	1	US-08-268-251-42	Sequence 42, Appl	651	6	0.8	29	4	US-08-834-130A-28	Sequence 28, Appl
579	6	0.8	23	1	US-08-268-251-52	Sequence 52, Appl	652	6	0.9	30	2	US-08-723-306-25	Sequence 25, Appl
580	6	0.9	23	1	US-08-393-985-24	Sequence 24, Appl	653	6	0.9	30	2	US-08-723-306-30	Sequence 30, Appl
581	6	0.9	23	2	US-08-652-450A-8	Sequence 8, Appl	654	6	0.9	30	4	US-09-348-578-5	Sequence 5, Appl
582	6	0.9	23	2	US-08-652-450A-10	Sequence 10, Appl	655	6	0.8	30	4	US-07-861-458B-132	Sequence 132, App
583	6	0.9	23	2	US-08-652-450A-14	Sequence 14, Appl	656	6	0.8	30	4	US-09-043-816B-14	Sequence 14, Appl
584	6	0.9	23	2	US-08-652-450A-15	Sequence 15, Appl	657	6	0.9	30	4	US-09-699-684-5	Sequence 5, Appl
585	6	0.9	23	2	US-08-652-450A-17	Sequence 17, Appl	658	6	0.9	30	5	PCT-US94-100461-31	Sequence 31, Appl
586	6	0.9	23	2	US-08-652-450A-18	Sequence 18, Appl	659	6	0.9	30	5	PCT-US96-100461-25	Sequence 25, Appl
587	6	0.9	23	2	US-08-612-999-2	Sequence 2, Appl	660	6	0.9	30	5	PCT-US96-10041-30	Sequence 30, Appl
588	6	0.9	23	3	US-08-701-382-2	Sequence 2, Appl	661	6	0.8	31	3	US-07-401-432-52	Sequence 52, Appl
589	6	0.8	23	3	US-08-256-747C-60	Sequence 60, Appl	662	6	0.9	31	4	US-09-348-578-14	Sequence 14, Appl
590	6	0.8	23	4	US-08-834-130A-60	Sequence 60, Appl	663	6	0.9	31	4	US-09-348-578-14	Sequence 14, Appl
591	6	0.8	23	4	US-08-557-006C-30	Sequence 30, Appl	664	6	0.8	31	4	US-09-023-731-8	Sequence 8, Appl
592	6	0.9	23	4	US-09-028-937-2	Sequence 2, Appl	665	6	0.8	31	4	US-09-327-357-491	Sequence 491, App
593	6	0.8	23	4	US-09-227-357-467	Sequence 467, App	666	6	0.9	31	4	US-09-699-684-6	Sequence 6, Appl
594	6	0.9	23	4	US-08-768-820-2	Sequence 2, Appl	667	6	0.9	31	4	US-09-699-684-14	Sequence 14, Appl
595	6	0.8	23	5	PCT-US93-01112-42	Sequence 42, Appl	668	6	0.8	32	1	US-07-952-735A-9	Sequence 9, Appl
596	6	0.8	23	5	PCT-US93-01112-52	Sequence 52, Appl	669	6	0.8	32	1	US-07-952-735A-10	Sequence 10, Appl

670	6	0.8	32	1	US-07-952-735A-11	Sequence 11, Appl	743	6	0.8	35	4	US-09-315-304B-1466	Sequence 1466, App
671	6	0.8	32	1	US-08-424-866-3	Sequence 3, Appl	C 744	6	0.9	35	4	US-09-699-684-18	Sequence 18, Appl
672	6	0.8	32	1	US-08-468-700-38	Sequence 38, Appl	C 745	6	0.9	35	4	US-09-699-684-26	Sequence 26, Appl
673	6	0.8	32	2	US-08-468-230-38	Sequence 38, Appl	746	6	0.8	35	6	5171673-7	Patent No. 5171673
674	6	0.8	32	2	US-08-468-688-38	Sequence 38, Appl	747	6	0.8	36	3	US-08-356-747C-25	Sequence 25, Appl
675	6	0.8	32	2	US-08-704-706A-38	Sequence 38, Appl	C 748	6	0.9	36	3	US-09-348-578-27	Sequence 27, Appl
676	6	0.8	32	2	US-08-595-868C-44	Sequence 44, Appl	749	6	0.8	36	4	US-08-834-130A-25	Sequence 25, Appl
677	6	0.8	32	2	US-08-595-868C-44	Sequence 44, Appl	C 750	6	0.9	36	4	US-09-699-684-27	Sequence 27, Appl
678	6	0.8	32	2	US-08-595-868C-46	Sequence 46, Appl	C 751	6	0.8	37	3	US-08-545-196B-64	Sequence 64, Appl
679	6	0.9	32	3	US-08-938-548B-7	Sequence 7, Appl	C 752	6	0.9	37	4	US-09-327-357-209	Sequence 209, App
680	6	0.8	32	3	US-07-401-422-53	Sequence 53, Appl	C 753	6	0.9	38	4	US-08-525-539A-2	Sequence 2, Appl
681	6	0.9	32	4	US-09-348-578-15	Sequence 7, Appl	C 754	6	0.9	38	4	US-09-336-536C-49	Sequence 49, Appl
682	6	0.9	32	4	US-09-348-578-15	Sequence 15, Appl	755	6	0.8	39	3	US-08-554-385-22	Sequence 22, Appl
683	6	0.8	32	4	US-09-348-578-23	Sequence 23, Appl	756	6	0.8	39	3	US-08-356-747C-64	Sequence 64, Appl
684	6	0.8	32	4	US-08-828-533-9	Sequence 9, Appl	757	6	0.8	39	4	US-08-834-130A-64	Sequence 64, Appl
685	6	0.8	32	4	US-08-985-659-39	Sequence 39, Appl	758	6	0.8	41	1	US-08-865-773-10	Sequence 10, Appl
686	6	0.8	32	4	US-09-139-819A-44	Sequence 44, Appl	759	6	0.8	41	2	US-08-117-952-794	Sequence 794, App
687	6	0.8	32	4	US-09-139-819A-45	Sequence 45, Appl	C 760	6	0.9	41	2	US-08-245-511-33	Sequence 33, Appl
688	6	0.8	32	4	US-09-139-819A-46	Sequence 46, Appl	C 761	6	0.9	41	2	US-08-600-993A-33	Sequence 33, Appl
689	6	0.8	32	4	US-08-194-664A-38	Sequence 38, Appl	762	6	0.8	41	4	US-08-981-189B-6	Sequence 6, Appl
690	6	0.9	32	4	US-08-939-093A-7	Sequence 7, Appl	763	6	0.8	41	4	US-09-400-716-8	Sequence 8, Appl
691	6	0.8	32	4	US-09-227-357-207	Sequence 207, App	764	6	0.8	41	4	US-09-424-127-18	Sequence 10, Appl
692	6	0.8	32	4	US-09-750-913-44	Sequence 44, Appl	765	6	0.8	43	3	US-08-256-747C-27	Sequence 27, Appl
693	6	0.8	32	4	US-09-750-913-45	Sequence 45, Appl	766	6	0.8	43	4	US-08-574-549A-202	Sequence 202, App
694	6	0.8	32	4	US-09-750-913-46	Sequence 46, Appl	767	6	0.8	43	4	US-08-834-130A-27	Sequence 27, Appl
695	6	0.9	32	4	US-09-699-684-7	Sequence 7, Appl	768	6	0.8	43	4	US-08-750-624-14	Sequence 14, Appl
696	6	0.9	32	4	US-09-699-684-15	Sequence 15, Appl	C 769	6	0.9	43	4	US-07-757-022B-12	Sequence 12, Appl
697	6	0.9	32	4	US-09-699-684-23	Sequence 23, Appl	C 770	6	0.9	44	1	US-08-362-037-133	Sequence 133, App
698	6	0.8	32	5	PCT-US94-01553A-38	Sequence 38, Appl	771	6	0.8	44	4	US-09-227-357-220	Sequence 220, App
699	6	0.8	32	5	PCT-US95-10426-38	Sequence 38, Appl	772	6	0.8	45	3	US-08-256-747C-65	Sequence 65, Appl
700	6	0.8	32	5	PCT-US96-05372-3	Sequence 3, Appl	773	6	0.9	46	4	US-08-483-533-30	Sequence 30, Appl
701	6	0.8	32	6	5183802-6	Patent No. 5183802	C 774	6	0.8	46	4	US-08-483-533-30	Sequence 30, Appl
702	6	0.8	32	6	5183802-7	Patent No. 5183802	C 775	6	0.9	46	4	US-09-283-471A-30	Sequence 30, Appl
703	6	0.8	32	6	5183802-8	Patent No. 5183802	C 776	6	0.9	46	4	US-07-757-022B-18	Sequence 18, Appl
704	6	0.8	33	1	US-07-776-272-5	Sequence 5, Appl	C 777	6	0.9	48	4	US-09-004-406C-15	Sequence 15, Appl
705	6	0.8	33	1	US-07-776-272-6	Sequence 6, Appl	778	6	0.8	48	4	US-09-288-143-102	Sequence 102, App
706	6	0.9	33	4	US-09-348-578-8	Sequence 8, Appl	779	6	0.8	50	1	US-08-321-071A-31	Sequence 31, Appl
707	6	0.9	33	4	US-09-348-578-16	Sequence 16, Appl	780	6	0.8	50	1	US-08-464-339A-16	Sequence 16, Appl
708	6	0.9	33	4	US-09-348-578-24	Sequence 24, Appl	C 781	6	0.9	50	3	US-08-630-916A-122	Sequence 122, App
709	6	0.9	33	4	US-09-699-684-8	Sequence 8, Appl	C 782	6	0.9	50	4	US-08-905-223-488	Sequence 488, App
710	6	0.9	33	4	US-09-699-684-16	Sequence 16, Appl	C 783	6	0.8	51	2	US-08-760-075A-27	Sequence 27, Appl
711	6	0.9	33	4	US-09-699-684-24	Sequence 24, Appl	C 784	6	0.9	51	4	US-08-444-818-2	Sequence 2, Appl
712	6	0.8	34	1	US-08-085-122-10	Sequence 10, Appl	C 785	6	0.8	51	4	US-09-338-546-27	Sequence 27, Appl
713	6	0.8	34	1	US-08-441-534A-13	Sequence 13, Appl	C 786	6	0.9	51	4	US-09-605-785-567	Sequence 567, App
714	6	0.8	34	1	US-08-629-752-13	Sequence 13, Appl	C 787	6	0.8	51	4	US-09-659-084-27	Sequence 27, Appl
715	6	0.9	34	1	US-08-623-195-1	Sequence 1, Appl	C 788	6	0.9	51	4	US-07-757-022B-100	Sequence 100, App
716	6	0.8	34	2	US-08-802-991-13	Sequence 13, Appl	C 789	6	0.9	51	4	US-07-757-022B-112	Sequence 112, App
717	6	0.9	34	2	US-08-902-516-14	Sequence 14, Appl	C 790	6	0.9	52	2	US-08-993-328-4	Sequence 4, Appl
718	6	0.9	34	2	US-08-319-052-20	Sequence 20, Appl	C 791	6	0.9	52	4	US-09-153-599A-4	Sequence 4, Appl
719	6	0.9	34	3	US-08-817-335-3	Sequence 3, Appl	C 792	6	0.9	53	1	US-08-222-616-6	Sequence 6, Appl
720	6	0.9	34	3	US-08-906-769-184	Sequence 184, App	C 793	6	0.9	53	4	US-08-446-648-6	Sequence 6, Appl
721	6	0.9	34	3	US-08-906-616-184	Sequence 184, App	C 794	6	0.9	53	5	PCT-US95-04228-6	Sequence 6, Appl
722	6	0.9	34	4	US-09-639-075A-184	Sequence 184, App	C 795	6	0.9	54	1	US-08-497-134A-16	Sequence 16, Appl
723	6	0.9	34	4	US-09-348-578-9	Sequence 9, Appl	C 796	6	0.9	54	1	US-08-464-339A-15	Sequence 15, Appl
724	6	0.9	34	4	US-09-348-578-17	Sequence 17, Appl	797	6	0.8	54	3	US-08-851-843A-19	Sequence 19, Appl
725	6	0.9	34	4	US-09-348-578-25	Sequence 25, Appl	798	6	0.8	54	4	US-08-854-050-19	Sequence 19, Appl
726	6	0.9	34	4	US-09-004-731-69	Sequence 69, Appl	799	6	0.8	54	4	US-09-430-323-19	Sequence 19, Appl
727	6	0.9	34	4	US-09-012-431-184	Sequence 69, App	C 800	6	0.9	55	4	US-09-361-707-91	Sequence 91, Appl
728	6	0.9	34	4	US-08-749-699-69	Sequence 69, Appl	C 801	6	0.9	55	4	US-09-361-707-93	Sequence 93, Appl
729	6	0.9	34	4	US-09-012-692-184	Sequence 184, App	C 802	6	0.9	55	4	US-09-361-707-95	Sequence 95, Appl
730	6	0.9	34	4	US-08-906-613-184	Sequence 184, App	C 803	6	0.9	55	4	US-09-361-707-96	Sequence 96, Appl
731	6	0.9	34	4	US-08-442-108B-20	Sequence 20, Appl	C 804	6	0.9	55	4	US-09-361-707-97	Sequence 97, Appl
732	6	0.9	34	4	US-09-397-680A-4	Sequence 4, Appl	C 805	6	0.9	55	4	US-09-361-707-99	Sequence 99, Appl
733	6	0.9	34	4	US-09-301-978C-19	Sequence 19, Appl	C 806	6	0.9	55	4	US-09-361-707-100	Sequence 100, App
734	6	0.9	34	4	US-09-004-729-69	Sequence 69, Appl	C 807	6	0.9	57	2	US-08-318-837-50	Sequence 50, Appl
735	6	0.9	34	4	US-09-699-684-9	Sequence 9, Appl	C 808	6	0.9	59	2	US-08-414-6570-52	Sequence 52, Appl
736	6	0.9	34	4	US-09-699-684-17	Sequence 17, Appl	C 809	6	0.9	59	2	US-08-414-6570-53	Sequence 53, Appl
737	6	0.9	34	4	US-09-699-684-25	Sequence 25, Appl	C 810	6	0.8	59	4	US-09-488-799-29	Sequence 29, Appl
738	6	0.9	35	2	US-08-652-450A-5	Sequence 5, Appl	C 811	6	0.9	60	4	US-09-327-357-227	Sequence 227, App
739	6	0.9	35	2	US-08-284-321B-47	Sequence 47, Appl	C 812	6	0.9	61	4	US-08-336-165A-408	Sequence 408, App
740	6	0.9	35	4	US-09-348-578-18	Sequence 18, Appl	C 813	6	0.8	62	4	US-09-227-357-231	Sequence 231, App
741	6	0.9	35	4	US-09-348-578-26	Sequence 26, Appl	C 814	6	0.8	64	1	US-07-689-693B-13	Sequence 13, Appl
742	6	0.9	35	4	US-09-218-950-47	Sequence 47, Appl	C 815	6	0.8	64	1	US-08-137-800-36	Sequence 36, Appl

816	6	0.8	64	1	US-08-137-800-45	Sequence 45, Appl	889	6	0.8	88	2	US-08-690-011A-31	Sequence 31, Appl
817	6	0.8	64	1	US-08-477-383-36	Sequence 36, Appl	890	6	0.8	88	3	US-09-045-632-7	Sequence 7, Appl
818	6	0.8	64	1	US-08-477-383-45	Sequence 45, Appl	891	6	0.9	88	4	US-09-330-970-6	Sequence 6, Appl
819	6	0.8	64	1	US-08-487-174-36	Sequence 36, Appl	892	6	0.9	88	4	US-09-263-933-23	Sequence 23, Appl
820	6	0.8	64	1	US-08-487-174-45	Sequence 45, Appl	893	6	0.9	89	4	US-09-214-631-13	Sequence 13, Appl
821	6	0.8	64	1	US-08-480-750-36	Sequence 36, Appl	894	6	0.9	91	2	US-08-997-080-143	Sequence 143, App
822	6	0.8	64	1	US-08-480-750-45	Sequence 45, Appl	895	6	0.9	91	2	US-08-997-362-143	Sequence 143, App
823	6	0.9	65	4	US-09-134-001C-3146	Sequence 3146, Ap	896	6	0.9	91	4	US-09-095-855-143	Sequence 143, App
824	6	0.9	67	4	US-09-003-198A-16	Sequence 16, Appl	897	6	0.9	91	4	US-09-324-542-143	Sequence 143, App
825	6	0.9	67	4	US-09-134-001C-5251	Sequence 5251, Ap	898	6	0.9	91	4	US-09-134-001C-5117	Sequence 5117, Ap
826	6	0.8	67	4	US-09-149-476-430	Sequence 430, App	899	6	0.9	91	4	US-09-205-426-143	Sequence 143, App
827	6	0.9	68	2	US-08-385-580A-20	Sequence 20, Appl	900	6	0.8	92	1	US-08-290-448A-68	Sequence 68, Appl
828	6	0.9	68	3	US-09-021-520-20	Sequence 20, Appl	901	6	0.8	92	1	US-08-290-448A-68	Sequence 68, Appl
829	6	0.9	68	4	US-09-134-001C-5173	Sequence 5173, Ap	902	6	0.8	92	1	US-08-175-069A-68	Sequence 68, Appl
830	6	0.9	69	1	US-08-358-160-107	Sequence 107, App	903	6	0.8	92	4	US-08-461-939B-68	Sequence 68, Appl
831	6	0.9	69	2	US-08-385-590A-18	Sequence 18, Appl	904	6	0.8	92	4	US-08-464-000-68	Sequence 68, Appl
832	6	0.9	69	3	US-09-021-520-18	Sequence 18, Appl	905	6	0.9	92	6	5166318-12	Patent No. 5166318
833	6	0.9	69	4	US-09-134-001C-3787	Sequence 3787, Ap	906	6	0.9	93	2	US-08-964-725-14	Sequence 14, Appl
834	6	0.8	70	4	US-08-905-223-290	Sequence 290, Appl	907	6	0.9	94	4	US-08-469-260A-170	Sequence 170, App
835	6	0.8	70	4	US-09-367-953B-70	Sequence 70, Appl	908	6	0.8	94	4	US-08-055-006-23	Sequence 23, Appl
836	6	0.8	70	4	US-09-367-953B-74	Sequence 74, Appl	909	6	0.9	95	4	US-09-188-930-344	Sequence 344, App
837	6	0.8	70	4	US-09-367-953B-78	Sequence 78, Appl	910	6	0.9	95	4	US-09-724-864-68	Sequence 68, Appl
838	6	0.8	70	4	US-09-367-953B-96	Sequence 96, Appl	911	6	0.8	95	4	US-09-134-001C-5541	Sequence 5541, Ap
839	6	0.9	70	4	US-09-006-428A-15	Sequence 15, Appl	912	6	0.8	96	3	US-09-101-167-10	Sequence 10, Appl
840	6	0.9	71	2	US-08-812-003-9	Sequence 9, Appl	913	6	0.8	97	2	US-08-242-097-5	Sequence 5, Appl
841	6	0.9	71	4	US-09-134-001C-2906	Sequence 2906, Ap	914	6	0.8	97	4	US-09-206-695-5	Sequence 5, Appl
842	6	0.8	71	4	US-08-469-260A-207	Sequence 207, App	915	6	0.8	97	4	US-09-309-487-23	Sequence 23, Appl
843	6	0.8	73	2	US-08-891-837B-2	Sequence 2, Appl	916	6	0.8	97	4	US-09-227-357-465	Sequence 45, App
844	6	0.8	73	4	US-09-134-001C-5316	Sequence 5316, Ap	917	6	0.9	97	4	US-09-422-936-73	Sequence 73, Appl
845	6	0.9	74	1	US-08-543-238-2	Sequence 2, Appl	918	6	0.9	97	4	US-09-422-936-83	Sequence 83, Appl
846	6	0.9	74	1	US-08-543-238-5	Sequence 5, Appl	919	6	0.9	98	2	US-08-825-556A-4	Sequence 4, Appl
847	6	0.9	74	1	US-08-420-526-2	Sequence 2, Appl	920	6	0.9	98	4	US-09-134-001C-4312	Sequence 4312, Ap
848	6	0.9	74	1	US-08-420-526-5	Sequence 5, Appl	921	6	0.9	99	2	US-08-825-556A-3	Sequence 3, Appl
849	6	0.8	74	1	US-08-594-031-97	Sequence 97, Appl	922	6	0.9	99	2	US-08-484-905-62	Sequence 62, Appl
850	6	0.8	74	3	US-08-866-545-6	Sequence 6, Appl	923	6	0.9	99	3	US-08-481-985B-62	Sequence 62, Appl
851	6	0.8	75	3	US-08-256-747C-56	Sequence 56, Appl	924	6	0.9	99	4	US-09-188-930-340	Sequence 340, App
852	6	0.8	75	4	US-08-834-130A-56	Sequence 56, Appl	925	6	0.9	99	4	US-08-370-476-62	Sequence 62, Appl
853	6	0.9	75	4	US-09-134-001C-5507	Sequence 5507, Ap	926	6	0.9	100	4	US-08-679-493A-146	Sequence 146, App
854	6	0.9	76	5	PCT-US95-13658-6	Sequence 6, Appl	927	6	0.8	100	2	US-08-808-982-8	Sequence 8, Appl
855	6	0.9	78	4	US-09-130-339-4	Sequence 4, Appl	928	6	0.9	102	2	US-09-174-060-2	Sequence 2, Appl
856	6	0.9	79	2	US-08-633-879C-14	Sequence 14, Appl	929	6	0.9	102	3	US-08-338-382-8	Sequence 8, Appl
857	6	0.9	79	2	US-09-174-060-1	Sequence 1, Appl	930	6	0.8	102	4	US-09-306-902A-8	Sequence 8, Appl
858	6	0.9	79	3	US-08-338-382-1	Sequence 1, Appl	931	6	0.9	102	4	US-09-134-001C-4995	Sequence 4995, Ap
859	6	0.9	79	4	US-09-100-802-4	Sequence 4, Appl	932	6	0.8	103	2	US-08-529-878B-41	Sequence 41, Appl
860	6	0.8	80	1	US-08-011-398B-14	Sequence 14, Appl	933	6	0.9	103	2	US-08-529-878B-41	Sequence 41, Appl
861	6	0.8	80	1	US-08-011-398B-15	Sequence 15, Appl	934	6	0.9	103	3	US-08-783-974-1	Sequence 1, Appl
862	6	0.8	80	1	US-08-011-398B-16	Sequence 16, Appl	935	6	0.9	105	1	US-08-467-420A-16	Sequence 16, Appl
863	6	0.8	80	1	US-08-464-051-14	Sequence 14, Appl	936	6	0.9	105	2	US-08-470-110A-16	Sequence 16, Appl
864	6	0.8	80	1	US-08-464-051-15	Sequence 15, Appl	937	6	0.9	105	2	US-08-940-371-16	Sequence 16, Appl
865	6	0.8	80	1	US-08-464-051-16	Sequence 16, Appl	938	6	0.9	105	3	US-08-488-551B-639	Sequence 639, App
866	6	0.8	80	2	US-08-462-498-14	Sequence 14, Appl	939	6	0.9	105	3	US-08-637-647-16	Sequence 16, Appl
867	6	0.8	80	2	US-08-462-498-15	Sequence 15, Appl	940	6	0.9	105	3	US-09-001-472-2	Sequence 2, Appl
868	6	0.8	80	2	US-08-462-498-16	Sequence 16, Appl	941	6	0.9	105	4	US-09-422-936-81	Sequence 81, Appl
869	6	0.9	80	3	US-08-462-498-16	Sequence 16, Appl	942	6	0.9	106	1	US-08-352-324A-5	Sequence 5, Appl
870	6	0.8	80	3	US-08-492-459-16	Sequence 16, Appl	943	6	0.9	106	1	US-08-862-607-5	Sequence 5, Appl
871	6	0.8	80	3	US-08-554-385-13	Sequence 13, Appl	944	6	0.9	106	2	US-08-468-819-7	Sequence 7, Appl
872	6	0.8	80	3	US-08-554-385-14	Sequence 14, Appl	945	6	0.9	106	2	US-08-468-819-7	Sequence 7, Appl
873	6	0.9	80	3	US-08-554-385-15	Sequence 15, Appl	946	6	0.9	106	3	US-08-352-324A-4	Sequence 4, Appl
874	6	0.9	80	4	US-08-423-752-16	Sequence 16, Appl	947	6	0.9	106	3	US-08-352-324A-4	Sequence 4, Appl
875	6	0.9	80	4	US-08-716-873-30	Sequence 30, Appl	948	6	0.9	106	4	US-08-662-607-7	Sequence 7, Appl
876	6	0.9	80	4	US-08-905-223-411	Sequence 411, Appl	949	6	0.9	106	4	US-08-662-607-7	Sequence 7, Appl
877	6	0.9	80	4	US-09-368-431-30	Sequence 30, Appl	950	6	0.9	106	4	US-09-189-129-4	Sequence 4, Appl
878	6	0.9	80	4	US-09-414-006-16	Sequence 16, Appl	951	6	0.9	106	5	PCT-US95-16144-5	Sequence 5, Appl
879	6	0.9	80	4	US-09-447-223-16	Sequence 16, Appl	952	6	0.9	107	1	US-08-161-939B-68	Sequence 68, Appl
880	6	0.9	81	2	US-08-248-839C-146	Sequence 146, App	953	6	0.9	107	1	US-08-352-324A-4	Sequence 4, Appl
881	6	0.9	81	2	US-08-245-511-14	Sequence 14, Appl	954	6	0.9	107	2	US-08-352-324A-4	Sequence 4, Appl
882	6	0.9	81	2	US-08-600-993A-14	Sequence 14, Appl	955	6	0.9	107	2	US-08-862-607-7	Sequence 7, Appl
883	6	0.9	82	1	US-09-100-802-5	Sequence 5, Appl	956	6	0.9	107	2	US-08-468-819-6	Sequence 6, Appl
884	6	0.8	82	4	US-08-497-312-30	Sequence 30, Appl	957	6	0.9	107	3	US-08-468-819-6	Sequence 6, Appl
885	6	0.8	83	4	US-09-513-783A-136	Sequence 136, App	958	6	0.9	107	3	US-09-203-235-4	Sequence 4, Appl
886	6	0.9	84	2	US-08-726-106A-76	Sequence 76, Appl	959	6	0.9	107	5	PCT-US95-16144-4	Sequence 4, Appl
887	6	0.9	84	4	US-09-386-493-8	Sequence 8, Appl	960	6	0.9	107	5	PCT-US95-16144-7	Sequence 7, Appl
888	6	0.8	87	4	US-09-299-495F-31	Sequence 31, Appl	961	6	0.9	108	1	US-08-468-661-3	Sequence 3, Appl

```
C 962 6 0.9 108 1 US-08-466-272A-3 Sequence 3, Appli
C 963 6 0.9 108 1 US-08-478-857-3 Sequence 3, Appli
C 964 6 0.9 108 1 US-08-488-113B-151 Sequence 151, App
C 965 6 0.9 108 1 US-08-477-484B-151 Sequence 151, App
C 966 6 0.9 108 1 US-08-107-669D-15 Sequence 15, Appl
C 967 6 0.9 108 1 US-08-472-788A-15 Sequence 15, Appl
C 968 6 0.9 108 2 US-08-477-531B-15 Sequence 15, Appl
C 969 6 0.9 108 2 US-08-646-360-151 Sequence 15, Appl
C 970 6 0.9 108 2 US-08-471-771-3 Sequence 3, Appli
C 971 6 0.9 108 2 US-08-082-842A-15 Sequence 15, Appl
C 972 6 0.9 108 3 US-08-388-353-539 Sequence 639, App
C 973 6 0.9 108 3 US-09-130-783-3 Sequence 3, Appli
C 974 6 0.8 108 3 US-08-959-212-10 Sequence 10, Appl
C 975 6 0.9 108 4 US-08-839-765-151 Sequence 151, App
C 976 6 0.9 108 4 US-09-136-389-151 Sequence 151, App
C 977 6 0.8 108 4 US-08-936-165A-443 Sequence 443, App
C 978 6 0.9 108 4 US-09-199-637A-227 Sequence 227, App
C 979 6 0.9 108 4 US-09-610-838-151 Sequence 151, App
C 980 6 0.8 109 2 US-08-210-762E-22 Sequence 22, Appl
C 981 6 0.9 109 2 US-08-308-494A-23 Sequence 23, Appl
C 982 6 0.8 109 4 US-09-106-075A-22 Sequence 22, Appl
C 983 6 0.8 110 4 US-09-134-001C-3110 Sequence 3110, Ap
C 984 6 0.9 110 4 US-09-134-001C-3110 Sequence 3110, Ap
C 985 6 0.8 110 4 US-09-134-001C-5023 Sequence 5023, Ap
C 986 6 0.8 110 6 5514582-9 Patent No. 5514582
C 987 6 0.9 111 1 US-08-466-886-19 Sequence 19, Appl
C 988 6 0.9 111 2 US-08-825-556A-2 Sequence 2, Appli
C 989 6 0.9 111 2 US-08-800-198-6 Sequence 6, Appli
C 990 6 0.9 111 3 US-09-296-595-6 Sequence 6, Appli
C 991 6 0.9 111 3 US-08-469-617-19 Sequence 19, Appl
C 992 6 0.8 111 4 US-09-347-801-8 Sequence 8, Appli
C 993 6 0.8 111 4 US-09-134-001C-3859 Sequence 3859, Ap
C 994 6 0.9 112 1 US-08-053-171-16 Sequence 16, Appl
C 995 6 0.9 112 1 US-08-331-398A-49 Sequence 49, Appl
C 996 6 0.9 112 1 US-08-477-877B-87 Sequence 87, Appl
C 997 6 0.9 112 1 US-08-477-877B-88 Sequence 88, Appli
C 998 6 0.8 112 1 US-08-446-919A-5 Sequence 5, Appli
C 999 6 0.9 112 2 US-08-472-281A-87 Sequence 87, Appli
C1000 6 0.9 112 2 US-08-472-281A-88 Sequence 88, Appli
```

## ALIGNMENTS

```
RESULT 1
US-09-058-489-15
; Sequence 15, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08PA
; CURRENT APPLICATION NUMBER: US/09/058, 489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041, 877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-15
```

```
Alignment Scores:
Pred. No.: 1,7e-12 Length: 662
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.04% Indels: 0
DB: 3 Gaps: 0
```

```
US-09-714-865-15 (1-2172) x US-09-058-489-15 (1-662)
QY 976 CGAGATTGATGGCTGTGCTCAACAGGCTGGAGACTGCGGCTTTCTCTACCA 1035
Db 218 ArgAspIeuMetAlaCysAlaGlnTrnGlySerGlyLysThrAlaIlePheLeuPro 237
QY 1036 ATTTTG 1041
Db 238 lleleu 239
```

```
RESULT 2
US-09-058-489-16
; Sequence 16, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08PA
; CURRENT APPLICATION NUMBER: US/09/058, 489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041, 877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-16

Alignment Scores:
Pred. No.: 1,7e-12 Length: 662
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.04% Indels: 0
DB: 3 Gaps: 0

US-09-714-865-15 (1-2172) x US-09-058-489-16 (1-662)
QY 976 CGAGATTGATGGCTGTGCTCAACAGGCTGGAGACTGCGGCTTTCTCTACCA 1035
Db 218 ArgAspIeuMetAlaCysAlaGlnTrnGlySerGlyLysThrAlaIlePheLeuPro 237
QY 1036 ATTTTG 1041
Db 238 lleleu 239

RESULT 3
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-08PA
; CURRENT APPLICATION NUMBER: US/09/058, 489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041, 877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
```

US-09-058-489-18

## Alignment Scores:

Pred. No.:	1.03e-08	Length:	660
Score:	18.00	Matches:	18
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.49%	Indels:	0
DB:	3	Gaps:	0

US-09-714-865-15 (1-2172) x US-09-058-489-18 (1-660)

QY 988 GCTTGCTCAACAGGCTGTGGAGACTGCGCTTTCTCTACCAATTTTG 1041

Db 220 AAcysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuProlIeLeu 237

## RESULT 4

US-09-058-489-91  
; Sequence 91, Application US/09058489  
; Patent No. 610386  
; GENERAL INFORMATION:  
; APPLICANT: Whitehead Institute for Biomedical Research  
; APPLICANT: Lahn, Bruce  
; TITLE OF INVENTION: Genes in the No. 610386-Recombining Region of  
; FILE REFERENCE: WH197-08PA  
; CURRENT APPLICATION NUMBER: US/09/058,489  
; CURRENT FILING DATE: 1998-04-10  
; EARLIER APPLICATION NUMBER: 60/041,877  
; EARLIER FILING DATE: 1997-04-11  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 91  
; LENGTH: 660  
; TYPE: PRT  
; ORGANISM: Human  
US-09-058-489-91

## Alignment Scores:

Pred. No.:	1.03e-08	Length:	660
Score:	18.00	Matches:	18
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.49%	Indels:	0
DB:	3	Gaps:	0

US-09-714-865-15 (1-2172) x US-09-058-489-91 (1-660)

QY 988 GCTTGCTCAACAGGCTGTGGAGACTGCGCTTTCTCTACCAATTTTG 1041

Db 220 AAcysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuProlIeLeu 237

## RESULT 5

US-09-183-706-43  
; Sequence 43, Application US/09183706  
; Patent No. 6245525  
; GENERAL INFORMATION:  
; APPLICANT: Martelange, Valérie  
; APPLICANT: De Smet, Charles  
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
; FILE REFERENCE: L0461/7054  
; CURRENT APPLICATION NUMBER: US/09/183,706  
; CURRENT FILING DATE: 1998-10-30  
; EARLIER APPLICATION NUMBER: 09/122,989  
; EARLIER FILING DATE: 1999-07-27  
; NUMBER OF SEQ ID NOS: 43  
; SEQ ID NO 43  
; LENGTH: 648  
; TYPE: PRT  
; ORGANISM: H. sapiens  
US-09-183-706-43

## Alignment Scores:

Pred. No.:	0.0432	Length:	648
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.52%	Indels:	0
DB:	4	Gaps:	0

US-09-714-865-15 (1-2172) x US-09-183-706-43 (1-648)

QY 1867 GAATATGTTTCATGCAATTGGCGCTACTGTCGT 1899

Db 570 GlutyrValHisArgIleGlyArgThrGlyArg 580

## RESULT 6

US-09-567-995-43  
; Sequence 43, Application US/09567995  
; Patent No. 6303756  
; GENERAL INFORMATION:  
; APPLICANT: Martelange, Valérie  
; APPLICANT: De Smet, Charles  
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
; FILE REFERENCE: L0461/7054  
; CURRENT APPLICATION NUMBER: US/09/567,995  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/183,706  
; PRIOR FILING DATE: 1998-10-30  
; NUMBER OF SEQ ID NOS: 43  
; SEQ ID NO 43  
; LENGTH: 648  
; TYPE: PRT  
; ORGANISM: H. sapiens  
US-09-567-995-43

## Alignment Scores:

Pred. No.:	0.0432	Length:	648
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.52%	Indels:	0
DB:	4	Gaps:	0

US-09-714-865-15 (1-2172) x US-09-567-995-43 (1-648)

QY 1867 GAATATGTTTCATGCAATTGGCGCTACTGTCGT 1899

Db 570 GlutyrValHisArgIleGlyArgThrGlyArg 580

## RESULT 7

US-08-679-493A-85  
; Sequence 85, Application US/08679493A  
; Patent No. 6303295  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Echan W.  
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS  
; FILE REFERENCE: 55-95  
; CURRENT APPLICATION NUMBER: US/08/679,493A  
; CURRENT FILING DATE: 1996-07-12  
; PRIOR APPLICATION NUMBER: 60/001203  
; PRIOR FILING DATE: 1995-07-14  
; PRIOR APPLICATION NUMBER: 60/003,112  
; PRIOR FILING DATE: 1995-09-01  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 85  
; LENGTH: 149  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-08-679-493A-85

## Alignment Scores:

OPERATING SYSTEM: PC-DOS/MS-DOS

Db 240 ILevalAlaProIlnrArgc

RESULT 11  
US-09-010-928B-2  
; Sequence 2, Application US/09010928B  
; Patent No. 5994099  
; GENERAL INFORMATION:  
; APPLICANT: Lewis, Randolph V  
; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
; STREET: 8110 GATEHOUSE RD. SUITE 500E  
; CITY: FALLS CHURCH  
; STATE: VIRGINIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 22042  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/010,928B  
; FILING DATE: 22-JAN-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M  
; REGISTRATION NUMBER: 28977  
; REFERENCE/DOCKET NUMBER: 1447-109P  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 870 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-010-928B-2

Alignment Scores:  
Pred. No.: 3.23 Length: 870  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.28% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-010-928B-2 (1-870)

QY 767 TCCTCATCTCAGGTGAGGAGGGGCT 741  
|||||  
Db 92 SerSerSerSerGlyGlyGlyGly 100

RESULT 12  
US-09-208-742-4  
; Sequence 4, Application US/09208742  
; Patent No. 6174679  
; GENERAL INFORMATION:  
; APPLICANT: Kaufmann, Joerg  
; TITLE OF INVENTION: CIFI50/HTAFI150 is Necessary for Cell  
; FILE REFERENCE: 1453.002  
; CURRENT APPLICATION NUMBER: US/09/208,742  
; CURRENT FILING DATE: 1998-12-10  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 1261  
; TYPE: PRT  
; ORGANISM: human  
US-09-208-742-4

Alignment Scores:  
Pred. No.: 3.06 Length: 1261

Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.24% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-208-742-4 (1-1261)

QY 1000 ACAGGCTCTGGAAGACTGGCGCTTTT 1026  
|||||  
Db 344 ThrGlySerGlySerThrAlaAlaIaphe 352

RESULT 13  
US-09-332-295-2  
; Sequence 2, Application US/09332295  
; Patent No. 6303372  
; GENERAL INFORMATION:  
; APPLICANT: Kaufmann, Joerg  
; TITLE OF INVENTION: CIFI30 INHIBITS CELL CYCLE PROGRESSION  
; FILE REFERENCE: 200130.456 / 1513.003  
; CURRENT APPLICATION NUMBER: US/09/332,295  
; CURRENT FILING DATE: 1999-06-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1261  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-332-295-2

Alignment Scores:  
Pred. No.: 3.06 Length: 1261  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.24% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-332-295-2 (1-1261)

QY 1000 ACAGGCTCTGGAAGACTGGCGCTTTT 1026  
|||||  
Db 344 ThrGlySerGlySerThrAlaAlaIaphe 352

RESULT 14  
US-09-709-979-2  
; Sequence 2, Application US/09709979  
; Patent No. 6423822  
; GENERAL INFORMATION:  
; APPLICANT: Kaufmann, Joerg  
; TITLE OF INVENTION: CIFI30 INHIBITS CELL CYCLE PROGRESSION  
; FILE REFERENCE: 200130.456 / 1513.003  
; CURRENT APPLICATION NUMBER: US/09/709,979  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: US 09/332,295  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1261  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-709-979-2

Alignment Scores:  
Pred. No.: 3.06 Length: 1261  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.24% Indels: 0  
Gaps: 0

US-09-714-865-15 (1-2172) x US-09-709-979-2 (1-1261)

QY 1000 ACAGGTCGGAAGACTGCGCTTTT 1026  
|||||  
Db 344 ThrGlySerGlyThrAlaAlaPhe 352

RESULT 15  
US-08-264-093-24

/ Sequence 24, Application US/08264093  
/ Patent No. 5639863  
/ GENERAL INFORMATION:  
/ APPLICANT: Michael D. Dan  
/ TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES SPECIFIC TO  
/ TITLE OF INVENTION: CELL CYCLE-INDEPENDENT GLIOMA SURFACE  
/ TITLE OF INVENTION: ANTIGEN  
/ NUMBER OF SEQUENCES: 26  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Rigout & Maybee  
/ STREET: 2300 Richmond-Adeleide Centre  
/ STREET: 101 Richmond Street West  
/ CITY: Toronto  
/ STATE: Ontario  
/ COUNTRY: Canada  
/ ZIP: M5H 2J7  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Diskette - 3.5 inch, 1.4 Mb storage  
/ COMPUTER: IBM PC Compatible  
/ OPERATING SYSTEM: MS-DOS 6.00  
/ SOFTWARE: ASCII Editor  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/264,093  
/ FILING DATE:  
/ CLASSIFICATION: 536  
/ PRIOR APPLICATION DATA: No. 5639863 applicable  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: Lake, James R.  
/ REGISTRATION NUMBER: 31081  
/ REFERENCE/DOCKET NUMBER: NOVOP/106A/7551  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (416) 868-1482  
/ TELEFAX: (416) 362-0823  
/ INFORMATION FOR SEQ ID NO: 24:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 17 amino acids  
/ TYPE: amino acid  
/ STRANDEDNESS: not applicable  
/ TOPOLOGY: linear  
/ US-08-264-093-24

## Alignment Scores:

Pred. NO.:	49.7	Length:	17
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.14%	Indels:	0
DB:	1	Gaps:	0

US-09-714-865-15 (1-2172) x US-08-264-093-24 (1-17)

QY 356 TCTTGCAGTCACTACTAGACTCT 333  
|||||  
Db 2 SerSerGlnSerLeuLeuAspSer 9

Search completed: June 10, 2003, 17:18:27  
Job time : 42.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:44:46 ; Search time 53.5 Seconds  
(without alignments)  
10819.452 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 724  
Sequence: 1 atgagggagatgaagatcggga.....tagatgatgagatcgggat 2172

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delcxt 7.0	

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters:  
-MODE=frame+ n2p.model -DEV=xid  
-O=/cgm2\_1/USPTO.spool/US0971465/runat.05062003\_111832\_26561/app\_query.fasta.1.2311  
-DB-A.geneseq.101002 -QMT=fastan -SUFFIX=oligo.rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=1000 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFM=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0971465 @CGN 1.1 76 @runat.05062003\_111832\_26561 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -NAIT -DSFBLCK=100 -LONGIOG  
-DEXTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A.Geneseq.101002.\*

1:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	724	100.0	724	22	AAE02417 Human vasa protein
2	123	17.0	135	22	AAU23109 Novel human enzyme
3	84	11.6	106	22	ABB95898 Human testicular
4	84	11.6	106	22	AAU23664 Novel human enzyme
5	84	11.6	106	22	AAW95199 Human reproductive
6	82	11.3	713	22	AAE02419 Rat vasa protein.
7	64	8.8	722	22	AAE02418 Mouse vasa protein
8	30	4.1	30	22	AAE02424 Epitope #2 of huma
9	29	4.0	29	22	AAE02424 Frog vasa protein.
10	24	3.3	700	22	AAE02420 Dantio reio vasa pr
11	24	3.3	700	22	AAE02421 Drosophila melanog
12	23	3.2	661	22	ABB59954 Fruit fly vasa pro
13	23	3.2	661	22	AAE02422 Gene 6 human secre
14	22	3.0	112	21	AAE02422 Human secreted pro
15	22	3.0	112	21	AAE02422 Dead Box X (DBX) g
16	22	3.0	662	19	AAW81502 Human DBX1, an RNA
17	22	3.0	662	19	AAW81501 Mouse ischaemic co
18	22	3.0	662	21	AAE02422 Arabidopsis thalia
19	22	2.6	646	21	AAE02422 Arabidopsis thalia
20	19	2.6	798	22	ABB64631 Arabidopsis thalia
21	19	2.6	798	22	AAE02422 Arabidopsis thalia
22	18	2.5	660	19	AAE02422 Arabidopsis thalia
23	17	2.3	612	21	AAE02422 Arabidopsis thalia
24	17	2.3	612	21	AAE02422 Arabidopsis thalia
25	16	2.2	344	21	AAE02422 Arabidopsis thalia
26	15	2.1	168	21	AAE02422 Arabidopsis thalia
27	15	2.1	257	21	AAE02422 Arabidopsis thalia
28	15	2.1	343	21	AAE02422 Arabidopsis thalia
29	15	2.1	500	22	AAE02422 Arabidopsis thalia
30	15	2.1	574	22	ABB67213 Arabidopsis thalia
31	15	2.1	577	22	ABB64981 Arabidopsis thalia
32	15	2.1	578	22	ABB65733 Arabidopsis thalia
33	15	2.1	578	22	ABB67212 Arabidopsis thalia
34	15	2.1	614	22	ABB44571 Arabidopsis thalia
35	15	2.1	614	22	ABB44572 Arabidopsis thalia
36	15	2.1	615	22	ABB44573 Arabidopsis thalia
37	15	2.1	615	22	ABB44573 Arabidopsis thalia
38	15	2.1	811	22	ABB64859 Arabidopsis thalia
39	15	2.1	945	22	ABB65231 Arabidopsis thalia
40	14	1.9	421	21	AAE02422 Arabidopsis thalia
41	14	1.9	421	21	AAE02422 Arabidopsis thalia
42	14	1.9	460	21	AAE02422 Arabidopsis thalia
43	14	1.9	722	22	ABG10453 Arabidopsis thalia
44	13	1.8	149	23	AAU99918 Arabidopsis thalia
45	13	1.8	457	23	AAU72977 Arabidopsis thalia
46	12	1.7	136	23	ABR07185 Arabidopsis thalia
47	12	1.7	222	22	AAU23649 Arabidopsis thalia
48	12	1.7	359	21	AAE02422 Arabidopsis thalia
49	12	1.7	359	21	AAE02422 Arabidopsis thalia
50	12	1.7	438	21	AAE02422 Arabidopsis thalia
51	12	1.7	619	22	AAE02422 Arabidopsis thalia
52	12	1.7	619	22	AAE02422 Arabidopsis thalia
53	12	1.7	619	22	AAE02422 Arabidopsis thalia
54	12	1.7	619	22	AAE02422 Arabidopsis thalia
55	11	1.5	45	23	ABG47079 Arabidopsis thalia
56	11	1.5	518	22	ABG25120 Arabidopsis thalia
57	11	1.5	589	19	AAW60667 Arabidopsis thalia
58	11	1.5	613	22	AAU3597 Arabidopsis thalia
59	11	1.5	613	22	AAU3597 Arabidopsis thalia
60	11	1.5	613	22	AAU3597 Arabidopsis thalia
61	11	1.5	613	22	AAU3597 Arabidopsis thalia
62	11	1.5	613	22	AAU3597 Arabidopsis thalia
63	11	1.5	613	22	AAU3597 Arabidopsis thalia
64	11	1.5	613	22	AAU3597 Arabidopsis thalia
65	10	1.4	1006	22	ABG24532 Arabidopsis thalia
66	10	1.4	1006	22	ABG24532 Arabidopsis thalia
67	10	1.4	1006	22	ABG24532 Arabidopsis thalia
68	10	1.4	1006	22	ABG24532 Arabidopsis thalia
69	10	1.4	1006	22	ABG24532 Arabidopsis thalia
70	10	1.4	1006	22	ABG24532 Arabidopsis thalia
71	10	1.4	1006	22	ABG24532 Arabidopsis thalia
72	10	1.4	1006	22	ABG24532 Arabidopsis thalia
73	10	1.4	1006	22	ABG24532 Arabidopsis thalia
74	10	1.4	1006	22	ABG24532 Arabidopsis thalia
75	10	1.4	1006	22	ABG24532 Arabidopsis thalia
76	10	1.4	1006	22	ABG24532 Arabidopsis thalia
77	10	1.4	1006	22	ABG24532 Arabidopsis thalia
78	10	1.4	1006	22	ABG24532 Arabidopsis thalia
79	10	1.4	1006	22	ABG24532 Arabidopsis thalia
80	10	1.4	1006	22	ABG24532 Arabidopsis thalia
81	10	1.4	1006	22	ABG24532 Arabidopsis thalia
82	10	1.4	1006	22	ABG24532 Arabidopsis thalia
83	10	1.4	1006	22	ABG24532 Arabidopsis thalia
84	10	1.4	1006	22	ABG24532 Arabidopsis thalia
85	10	1.4	1006	22	ABG24532 Arabidopsis thalia
86	10	1.4	1006	22	ABG24532 Arabidopsis thalia
87	10	1.4	1006	22	ABG24532 Arabidopsis thalia
88	10	1.4	1006	22	ABG24532 Arabidopsis thalia
89	10	1.4	1006	22	ABG24532 Arabidopsis thalia
90	10	1.4	1006	22	ABG24532 Arabidopsis thalia
91	10	1.4	1006	22	ABG24532 Arabidopsis thalia
92	10	1.4	1006	22	ABG24532 Arabidopsis thalia
93	10	1.4	1006	22	ABG24532 Arabidopsis thalia
94	10	1.4	1006	22	ABG24532 Arabidopsis thalia
95	10	1.4	1006	22	ABG24532 Arabidopsis thalia
96	10	1.4	1006	22	ABG24532 Arabidopsis thalia
97	10	1.4	1006	22	ABG24532 Arabidopsis thalia
98	10	1.4	1006	22	ABG24532 Arabidopsis thalia
99	10	1.4	1006	22	ABG24532 Arabidopsis thalia
100	10	1.4	1006	22	ABG24532 Arabidopsis thalia

69	10	1.4	221	22	AU23651	Novel human enzyme	142	9	1.2	768	21	AA842448	Human ORFX ORF2212
70	10	1.4	226	22	AU23095	Novel human enzyme	143	9	1.2	782	22	AB58923	Drosophila melanog
71	10	1.4	303	21	AA646347	Arabidopsis thaliana	144	9	1.2	791	22	AB862354	Drosophila melanog
72	10	1.4	323	22	ABG10452	Novel human diagno	145	9	1.3	924	21	AA522991	Drosophila melanog
73	10	1.4	327	21	AA646346	Arabidopsis thaliana	146	9	1.3	958	22	AB862764	Drosophila melanog
74	10	1.4	421	22	AAU36110	Klebsiella pneumonia	147	9	1.2	1275	20	AA429084	Human C1F130 prote
75	10	1.4	424	22	AAU36065	C glutathion prote	148	9	1.2	1275	21	AA474262	Human cell cycle r
76	10	1.4	440	22	AU23549	Novel human enzyme	149	9	1.3	3502	22	AB583882	Drosophila melanog
77	10	1.4	456	21	AA646345	Arabidopsis thaliana	150	8	1.1	16	13	AA824704	Sequence of an ant
78	10	1.4	460	22	ABG08818	Novel human diagno	151	8	1.1	16	13	AA866292	Murine derived ant
79	10	1.4	470	23	AB848320	Listeria monocytog	152	8	1.1	16	22	AB873652	Murine monoclonal
80	10	1.4	484	23	AA825668	Human protein sequ	153	8	1.1	16	23	AAO14984	Mouse mab ID9 anti
81	10	1.4	495	23	AAE20093	Lactobacillus rham	154	8	1.1	20	21	AA470796	Murine anti-p53 PA
82	10	1.4	497	20	AA41671	Lactobacillus reut	155	8	1.1	25	17	AA888856	Peptide targeted
83	10	1.4	520	23	AB847602	Listeria monocytog	156	8	1.1	39	22	AB844289	Peptide #11795 enc
84	10	1.4	524	22	AAU37898	Streptococcus pneu	157	8	1.1	39	22	AB827151	Protein #9150 enco
85	10	1.4	528	23	ABP28778	Streptococcus poly	158	8	1.1	39	22	AA865326	Human brain expres
86	10	1.4	528	23	ABP29908	Streptococcus poly	159	8	1.1	39	22	AA878023	Human bone marrow
87	10	1.4	536	22	AU35181	Enterococcus faeca	160	8	1.1	39	22	AA421912	Peptide #8346 enco
88	10	1.4	539	23	ABP28779	Streptococcus poly	161	8	1.1	39	22	AA838240	Peptide #12277 enc
89	10	1.4	547	23	AB853657	Lactococcus lactis	162	8	1.1	39	23	ABG47038	Human peptide enco
90	10	1.4	567	22	AU36292	Pseudomonas aerugi	163	8	1.1	40	21	AA656574	Arabidopsis thalia
91	10	1.4	568	22	AU36294	Propionibacterium	164	8	1.1	53	23	ABP09771	Human ORFX protein
92	10	1.4	619	22	AB869630	Drosophila melanog	165	8	1.1	58	22	AAU22441	Human cardiovascular
93	10	1.4	622	22	AB892773	Human protein sequ	166	8	1.1	61	23	ABP34492	Human helicase-11k
94	10	1.4	622	22	AA83391	Human protein sequ	167	8	1.1	63	21	AA650020	Arabidopsis thalia
95	10	1.4	624	23	AA814675	Corynebacterium gl	168	8	1.1	64	22	AA8966330	Human reproductive
96	10	1.4	628	21	AA842879	Human ORFX ORF643	169	8	1.1	65	22	AA806085	Human gene 45 enco
97	10	1.4	635	20	AA73847	Human prostate tum	170	8	1.1	65	23	ABG33907	Human secreted pro
98	10	1.4	648	22	ABG22292	Novel human diagno	171	8	1.1	66	21	AA878708	Human secreted pro
99	10	1.4	732	22	AA891024	C glutathion prote	172	8	1.1	69	23	ABP09715	Human ORFX protein
100	10	1.4	1038	22	ABG24868	Novel human diagno	173	8	1.1	72	20	AA706406	Human secreted pro
101	10	1.4	1224	22	AB862191	Drosophila melanog	174	8	1.1	72	23	AAU39017	Human secreted pro
102	9	1.3	40	21	AA838560	Human secreted pro	175	8	1.1	72	23	AB855726	Human polypeptide
103	9	1.2	59	22	AU23800	Novel human enzyme	176	8	1.1	73	23	ABP05652	Human ORFX protein
104	9	1.2	59	22	AA863264	Human immune/haema	177	8	1.1	75	22	AAU30402	Novel human secret
105	9	1.3	108	21	AA834179	Zea mays protein f	178	8	1.1	75	22	AAO10920	Human polypeptide
106	9	1.3	126	21	AA565568	Flagellin from fibro	179	8	1.1	78	16	AA870161	Streptococcus pneu
107	9	1.3	191	21	AA825522	Human adenovirus t	180	8	1.1	93	23	AAU91110	Human secreted pro
108	9	1.2	257	22	AA892816	Human protein sequ	181	8	1.1	95	21	AA877195	Human secreted pro
109	9	1.3	262	22	AB866342	Drosophila melanog	182	8	1.1	95	22	AAE06172	Human gene 45 enco
110	9	1.3	287	22	AB866343	Drosophila melanog	183	8	1.1	95	23	ABG33966	Human secreted pro
111	9	1.3	312	23	AB804728	Human PP3241 prote	184	8	1.1	99	21	AA609859	Arabidopsis thalia
112	9	1.3	325	22	AA848515	Fiber protein of A	185	8	1.1	99	21	AA625224	Arabidopsis thalia
113	9	1.3	337	21	AA568657	Amino acid sequenc	186	8	1.1	99	21	AA650019	Arabidopsis thalia
114	9	1.3	338	21	AA568656	Amino acid sequenc	187	8	1.1	99	21	AA606858	Murine ID9 antibod
115	9	1.3	338	22	AA831741	Amino acid sequenc	188	8	1.1	100	22	AAE06959	Mouse germline Kap
116	9	1.3	338	22	AA831742	Amino acid sequenc	189	8	1.1	110	21	AA770788	Murine anti-p53 mo
117	9	1.3	338	22	AA867184	Mastadenovirus 34	190	8	1.1	110	22	AA877590	Human bone marrow
118	9	1.3	338	22	AA867185	Mastadenovirus 35	191	8	1.1	110	23	AAO14974	Humanised murine 1
119	9	1.3	342	23	AAU79745	Amino acid sequenc	192	8	1.1	110	23	ABG46625	Human peptide enco
120	9	1.2	419	23	AU23554	Novel human enzyme	193	8	1.1	111	23	AAO14975	Humanised murine 1
121	9	1.2	441	23	AB848633	Listeria monocytog	194	8	1.1	111	23	AAO14977	Humanised murine 1
122	9	1.3	448	22	AB858601	Drosophila melanog	195	8	1.1	112	13	AA824710	Sequence of a chim
123	9	1.2	486	22	AB866093	Drosophila melanog	196	8	1.1	112	19	AA848248	A77 anti-Fc alpha
124	9	1.3	505	22	AA861923	Human early growth	197	8	1.1	112	20	AA522765	Anti-tissue factor
125	9	1.3	543	13	AA824022	Human promyeloleu	198	8	1.1	112	20	AA52766	Anti-tissue factor
126	9	1.3	543	20	AA876985	Human EGR-1 protei	199	8	1.1	112	22	AA893177	Anti-p53 monoclon
127	9	1.3	543	21	AA51116	Human EGR-1 protei	200	8	1.1	112	22	AAU09918	Murine mab ID9 11g
128	9	1.3	543	22	AAU09066	Human ZIF268 prote	201	8	1.1	112	22	AAU09921	Humanised ID9 11g
129	9	1.3	543	22	AAU76861	Human lung tumour	202	8	1.1	112	22	AAU09922	Humanised ID9 11g
130	9	1.3	543	23	AAU85516	Clone #18996 of lu	203	8	1.1	112	22	AAU09923	Humanised ID9 11g
131	9	1.3	543	22	AB859192	Arabidopsis thaliana	204	8	1.1	112	22	AAU09924	Humanised ID9 11g
132	9	1.2	556	22	AA873639	Arabidopsis thaliana	205	8	1.1	112	22	AAU09925	Humanised ID9 11g
133	9	1.3	574	23	AB897289	Novel human protei	206	8	1.1	112	22	AAE08554	Murine A77 anti-Fc
134	9	1.3	663	18	AA829150	Drosophila melanog	207	8	1.1	112	22	AAE06946	Murine ID9 antibod
135	9	1.2	703	22	ABG19359	Dual-specific muri	208	8	1.1	112	22	AAE06949	Humanised murine 1
136	9	1.2	709	22	AA876859	Novel human diagno	209	8	1.1	112	22	AAE06950	Humanised murine 1
137	9	1.2	709	23	AAU85514	Human lung tumour	210	8	1.1	112	22	AAE06951	Humanised murine 1
138	9	1.2	709	23	AAU85514	Human lung tumour	211	8	1.1	112	22	AAE06952	Humanised murine 1
139	9	1.2	746	23	AA837465	ATP-dependent RNA	212	8	1.1	112	22	AAE06991	Human kappa light
140	9	1.2	765	22	AA873637	Human ATP-dependen	213	8	1.1	112	22	AAE07036	Humanised murine 1
141	9	1.2	765	22	AA894373	Human protein sequ	214	8	1.1	112	22	AA874621	A77 anti-Fc-alpha

C 215	8	1.1	112	23	AA014971	Mouse light chain	C 288	8	1.1	312	21	AA642946	Arabidopsis thalia
C 216	8	1.1	112	23	AA014973	Humanised murine 1	C 289	8	1.1	314	22	ABG22291	Novel human diagno
C 217	8	1.1	112	23	AA014976	Humanised murine 1	C 290	8	1.1	320	21	AA617259	Arabidopsis thalia
C 218	8	1.1	113	17	AA888508	Vlkappa for antibo	C 291	8	1.1	320	21	AA642945	Arabidopsis thalia
C 219	8	1.1	113	22	AAE07005	Murine kappa light	C 292	8	1.1	320	22	AAE02556	A. thaliana transc
C 220	8	1.1	113	22	AAE03751	Murine PSKA Ab 1lg	C 293	8	1.1	320	22	AAU93119	Arabidopsis thali
C 221	8	1.1	114	22	AAE07035	Humanised murine a	C 294	8	1.1	325	22	AA848516	Fiber protein of A
C 222	8	1.1	116	22	ABG30374	Novel human diagno	C 295	8	1.1	328	22	AAU37963	Streptococcus pneu
C 223	8	1.1	118	22	ABG16698	Novel human diagno	C 296	8	1.1	328	22	AAU37963	Streptococcus pneu
C 224	8	1.1	119	23	AAU91087	Human secreted pro	C 297	8	1.1	328	22	ABG07252	Novel human diagno
C 225	8	1.1	123	23	ABG04683	Novel human diagno	C 298	8	1.1	326	21	AA643669	Arabidopsis thalia
C 226	8	1.1	126	23	AAU91137	Human secreted pro	C 299	8	1.1	326	21	AAU11818	Cancer and neuroge
C 227	8	1.1	131	12	AA812239	Mouse MAb 4D12 L c	C 300	8	1.1	349	21	AA614415	Arabidopsis thalia
C 228	8	1.1	131	12	AA010924	Human polypeptide	C 301	8	1.1	352	22	ABG07250	Novel human diagno
C 229	8	1.1	132	12	AA812361	Light (kappa) chai	C 302	8	1.1	353	20	AAU06273	Anti Fc alpha rece
C 230	8	1.1	132	13	AA824712	Sequence encoded b	C 303	8	1.1	353	21	AAU31755	Amino acid sequenc
C 231	8	1.1	133	21	AA843731	Human cancer assoc	C 304	8	1.1	353	22	AAU31755	Amino acid sequenc
C 232	8	1.1	138	22	ABG25114	Novel human diagno	C 305	8	1.1	353	22	AA867198	Mastadenovirus 51
C 233	8	1.1	142	22	AAE07032	Murine antibody 1D	C 306	8	1.1	356	21	AA635470	Arabidopsis thalia
C 234	8	1.1	143	19	AAW98594	H. pylori GHP 115	C 307	8	1.1	357	23	AAU79746	Amino acid sequenc
C 235	8	1.1	143	22	AA01800	Human polypeptide	C 308	8	1.1	357	23	AA643944	Arabidopsis thalia
C 236	8	1.1	145	22	AB868802	Drosophila melanog	C 309	8	1.1	361	21	AA643943	Zea mays protein f
C 237	8	1.1	152	21	AA609858	Arabidopsis thalia	C 310	8	1.1	363	21	AA643942	Zea mays protein f
C 238	8	1.1	158	22	ABG30370	Novel human diagno	C 311	8	1.1	367	23	AB810101	Human homeobox pro
C 239	8	1.1	160	21	AA609857	Arabidopsis thalia	C 312	8	1.1	368	23	AAU11273	Murine beta1,3-N-a
C 240	8	1.1	160	22	AA835101	Antibody variable	C 313	8	1.1	369	21	AAU11273	Arabidopsis thalia
C 241	8	1.1	161	22	AAW24370	Human EST encoded	C 314	8	1.1	371	23	AAU11819	Cancer and neuroge
C 242	8	1.1	166	21	AA617261	Arabidopsis thalia	C 315	8	1.1	374	22	AAW52309	Murine zyxine frag
C 243	8	1.1	166	21	AA625223	Arabidopsis thalia	C 316	8	1.1	380	22	AAW52311	Human zyxine fragm
C 244	8	1.1	166	21	AA642947	Arabidopsis thalia	C 317	8	1.1	387	21	AAW56899	Human prostate can
C 245	8	1.1	166	23	ABP39434	Staphylococcus epi	C 318	8	1.1	391	22	AAW5280	Human protein sequ
C 246	8	1.1	168	22	ABG07982	Novel human diagno	C 319	8	1.1	401	22	ABG08143	Novel human diagno
C 247	8	1.1	168	22	ABG30371	Novel human diagno	C 320	8	1.1	407	22	AA893496	Human protein sequ
C 248	8	1.1	187	20	AAV60035	Human endometrium	C 321	8	1.1	413	22	AA892764	Human protein sequ
C 249	8	1.1	189	21	AA625222	Arabidopsis thalia	C 322	8	1.1	413	22	AA894038	Human protein sequ
C 250	8	1.1	191	21	AAV82520	Human adenovirus t	C 323	8	1.1	418	21	AA641235	Arabidopsis thalia
C 251	8	1.1	203	21	AA621890	Arabidopsis thalia	C 324	8	1.1	439	21	AA641234	Arabidopsis thalia
C 252	8	1.1	204	21	AA619655	Arabidopsis thalia	C 325	8	1.1	440	22	AA692622	AA692622
C 253	8	1.1	205	21	AA621889	Arabidopsis thalia	C 326	8	1.1	440	22	AA879004	C. glutamicum prote
C 254	8	1.1	205	22	ABG07253	Novel human diagno	C 327	8	1.1	442	22	AA894926	Human protein sequ
C 255	8	1.1	206	21	AA619964	Arabidopsis thalia	C 328	8	1.1	443	22	AAW40069	Human protein sequ
C 256	8	1.1	207	21	AA619963	Arabidopsis thalia	C 329	8	1.1	447	23	ABP41581	Human ovairtan anti
C 257	8	1.1	210	21	AA621888	Arabidopsis thalia	C 330	8	1.1	457	21	AA654360	Arabidopsis thalia
C 258	8	1.1	211	21	AA654371	Human pancreatic c	C 331	8	1.1	459	21	AA643582	Human cancer assoc
C 259	8	1.1	213	21	AA604742	Arabidopsis thalia	C 332	8	1.1	461	21	AA642956	Human ORFX ORF2720
C 260	8	1.1	214	22	AAU23539	Novel human enzyme	C 333	8	1.1	461	21	AA654359	Arabidopsis thalia
C 261	8	1.1	216	21	AA604741	Arabidopsis thalia	C 334	8	1.1	462	22	AAU30404	Novel human secrec
C 262	8	1.1	218	22	AAE03756	Chimeric antibody	C 335	8	1.1	468	21	AA653119	Arabidopsis thalia
C 263	8	1.1	220	21	AA653303	Human colon cancer	C 336	8	1.1	471	23	ABG65636	Human breast speci
C 264	8	1.1	220	21	AA631903	Arabidopsis thalia	C 337	8	1.1	473	22	AB864634	Drosophila melanog
C 265	8	1.1	220	22	AA673751	Human colon cancer	C 338	8	1.1	478	22	AB893520	Human protein sequ
C 266	8	1.1	223	21	AA614416	Arabidopsis thalia	C 339	8	1.1	478	23	AB857344	Mouse ischemic co
C 267	8	1.1	224	20	AAV60124	Human endometrium	C 340	8	1.1	479	21	AA647287	Arabidopsis thalia
C 268	8	1.1	233	22	AAU27773	Human full-length	C 341	8	1.1	479	21	AAV58616	Protein regulating
C 269	8	1.1	233	22	AA895172	Human protein sequ	C 342	8	1.1	479	22	AAW39539	Human polypeptide
C 270	8	1.1	233	22	AA836593	Human FLEXHT-15 pr	C 343	8	1.1	479	22	AA893928	Human protein sequ
C 271	8	1.1	236	21	AA856832	Human prostate can	C 344	8	1.1	479	22	AA837464	Murine ATP-depende
C 272	8	1.1	243	19	AAW60769	Single chain anticb	C 345	8	1.1	483	23	AAW47315	Human ATP dependen
C 273	8	1.1	248	22	ABG08222	Novel human diagno	C 346	8	1.1	485	21	AA653118	Arabidopsis thalia
C 274	8	1.1	259	23	AA604740	Arabidopsis thalia	C 347	8	1.1	487	21	AA648622	Arabidopsis thalia
C 275	8	1.1	261	23	ABP43023	Human ovarian anti	C 348	8	1.1	487	22	AAU36804	Staphylococcus aur
C 276	8	1.1	271	18	AAU12692	Natural killer cel	C 349	8	1.1	487	22	AAU37398	Staphylococcus aur
C 277	8	1.1	271	21	AAV43399	Human natural kill	C 350	8	1.1	489	21	AA653117	Arabidopsis thalia
C 278	8	1.1	271	22	AAW50249	Human natural kill	C 351	8	1.1	491	22	AA893174	Human protein sequ
C 279	8	1.1	271	22	AA882478	Natural killer cel	C 352	8	1.1	492	22	AAU35715	Helicobacter pylor
C 280	8	1.1	271	22	AA888038	Amino acid sequenc	C 353	8	1.1	493	22	ABG07251	Novel human diagno
C 281	8	1.1	275	21	AA635471	Arabidopsis thalia	C 354	8	1.1	494	21	AA654358	Arabidopsis thalia
C 282	8	1.1	283	21	AA631902	Arabidopsis thalia	C 355	8	1.1	499	21	AA635469	Arabidopsis thalia
C 283	8	1.1	296	20	AAV41421	Raly protein sequ	C 356	8	1.1	502	21	AA611871	Arabidopsis thalia
C 284	8	1.1	302	21	AA643670	Arabidopsis thalia	C 357	8	1.1	502	21	AA646881	Arabidopsis thalia
C 285	8	1.1	304	21	AA631901	Arabidopsis thalia	C 358	8	1.1	510	22	AAU33760	Staphylococcus aur
C 286	8	1.1	307	22	AB858806	Drosophila melanog	C 359	8	1.1	511	22	AA67687	Amino acid sequenc
C 287	8	1.1	312	21	AA617260	Arabidopsis thalia	C 360	8	1.1	514	22	AAW41325	Human polypeptide

361	8	1.1	521	22	ABBS5234	Drosophila melanog	C 434	8	1.1	797	20	AAV05848	Banana ripening fr
C 362	8	1.1	524	22	ABBS4765	Drosophila melanog	435	8	1.1	801	22	AAW40858	Human polypeptide
C 363	8	1.1	526	22	AAW78355	Human protein SEQ	436	8	1.1	802	22	ABBS1892	Drosophila melanog
364	8	1.1	526	22	AAU00864	S. aureus ATP-depe	437	8	1.1	805	22	ABBS1739	Drosophila melanog
365	8	1.1	527	20	AAI05466	S. aureus cDPA pro	C 438	8	1.1	805	22	ABBS62304	Drosophila melanog
366	8	1.1	528	23	ABP39417	Staphylococcus epi	439	8	1.1	813	22	ABBS64229	Drosophila melanog
367	8	1.1	529	22	AAU33292	Novel human secret	440	8	1.1	859	20	AAV00991	Human ATP-dependen
C 368	8	1.1	535	18	AAW28491	Human p53 protein	441	8	1.1	859	22	AAW38810	Human polypeptide
C 369	8	1.1	535	18	AAW28492	Human p53 protein	442	8	1.1	859	22	AAW33268	Human protein HP02
C 370	8	1.1	536	22	AAW78356	Human protein SEQ	443	8	1.1	859	22	AAW37463	Human ATP-dependen
371	8	1.1	547	21	AAI56849	Human RNA binding	444	8	1.1	872	22	ABG07781	Novel human diagno
372	8	1.1	547	22	AAW39379	Human polypeptide	445	8	1.1	887	22	AAW40536	Human polypeptide
373	8	1.1	547	22	AAW39379	Human polypeptide	446	8	1.1	892	22	ABG30372	Novel human diagno
374	8	1.1	547	22	AAW39379	Human polypeptide	C 447	8	1.1	926	23	ABBS7572	Novel human diagno
375	8	1.1	548	20	AAW88549	Amino acid sequenc	C 448	8	1.1	951	22	AAE06694	Arabidopsis thalia
376	8	1.1	548	22	ABBS0316	Secretd protein e	C 449	8	1.1	965	21	AAW53461	Arabidopsis thalia
377	8	1.1	550	21	AAW48621	Human secreted pro	450	8	1.1	971	21	AAW53460	Arabidopsis thalia
C 378	8	1.1	552	22	ABBS1493	Arabidopsis thalia	451	8	1.1	1089	21	AAW53459	Arabidopsis thalia
C 379	8	1.1	552	22	ABBS1493	Drosophila melanog	C 452	8	1.1	1096	22	ABBS60222	Drosophila thalia
C 380	8	1.1	564	22	AAW52303	Drosophila melanog	C 453	8	1.1	1108	22	ABBS63413	Drosophila melanog
381	8	1.1	564	23	AAW52303	Mouse ischaemic co	C 454	8	1.1	1164	22	ABBS7802	Drosophila melanog
382	8	1.1	566	22	AAW41165	Human polypeptide	C 455	8	1.1	1338	22	ABBS70361	Drosophila melanog
C 383	8	1.1	569	22	AAW79339	Human protein SEQ	C 456	8	1.1	1368	22	ABBS65408	Drosophila melanog
C 384	8	1.1	569	22	AAW79339	Human protein SEQ	C 457	8	1.1	1783	22	ABBS63930	Drosophila melanog
385	8	1.1	571	21	AAW48620	Arabidopsis thalia	C 458	8	1.1	2090	22	ABBS64682	Drosophila melanog
386	8	1.1	572	22	AAW48620	Zyxin protein SEQ	C 459	8	1.1	2153	22	AAW33195	Novel human secret
387	8	1.1	572	22	AAW52305	Human protein SEQ	C 460	8	1.1	2175	22	ABBS65698	Drosophila melanog
388	8	1.1	574	22	AAW41855	Human polypeptide	C 461	8	1.1	2306	23	AAU11817	Cancer and neuroge
C 389	8	1.1	588	21	AAW39378	Arabidopsis thalia	C 462	8	1.1	2352	23	AAU11816	Cancer and neuroge
390	8	1.1	590	22	ABG22143	Novel human diagno	463	8	1.1	10	22	AAW5019	Human complementar
391	8	1.1	593	21	AAW39378	Arabidopsis thalia	464	8	1.1	10	22	AAW67833	Saccharomyces cere
392	8	1.1	593	21	AAW39378	Arabidopsis thalia	C 465	8	1.1	14	20	AAW67833	Human secreted pro
393	8	1.1	594	22	AAW42865	Arabidopsis thalia	C 466	8	1.1	17	21	AAW95231	Human secreted pro
394	8	1.1	594	22	ABBS61659	Drosophila melanog	467	8	1.1	18	17	AAW05474	SH3-binding antiody
395	8	1.1	599	23	AAE15278	Human RNA metaboli	468	8	1.1	18	22	ABBS40992	Peptide #8938 enco
396	8	1.1	600	23	AAU99903	Human 49875 DEAD t	469	8	1.1	18	22	ABBS25090	Protein #7089 enco
C 397	8	1.1	603	22	ABBS68633	Drosophila melanog	470	8	1.1	18	22	AAW61852	Human brain expres
398	8	1.1	610	22	ABG19757	Novel human diagno	471	8	1.1	18	22	AAW44652	Human bone marrow
399	8	1.1	612	21	AAW33869	Arabidopsis thalia	472	8	1.1	18	22	AAW34767	Peptide #8804 enco
400	8	1.1	612	21	AAW33869	Arabidopsis thalia	473	8	1.1	18	22	AAW61207	Human INTERCEPT 29
401	8	1.1	613	22	ABBS64879	Arabidopsis thalia	474	8	1.1	18	23	ABG44475	Human peptide enco
402	8	1.1	615	22	ABBS64540	Drosophila melanog	C 475	8	1.1	19	20	AAW11541	Human 5' EST seque
403	8	1.1	615	22	ABG07784	Novel human diagno	C 476	8	1.1	20	21	AAW45163	Human secreted pro
C 404	8	1.1	619	21	AAW45163	Novel human diagno	C 477	8	1.1	20	21	AAW45163	TANCO 300 signal p
C 405	8	1.1	627	20	AAW05851	Banana ripening fr	C 478	8	1.1	21	21	AAW34122	Human secreted pro
406	8	1.1	633	20	AAW44017	Saccharomyces cere	C 479	8	1.1	22	21	AAW33186	Arab promoter and
407	8	1.1	637	21	AAW17814	Arabidopsis thalia	C 480	8	1.1	22	20	AAW36382	Fragment of human
408	8	1.1	657	21	AAW49207	Arabidopsis thalia	481	8	1.1	22	20	AAW01199	Polypeptide fragme
409	8	1.1	663	22	ABBS62293	Drosophila melanog	C 482	8	1.1	28	22	ABBS27725	Human peptide #376
410	8	1.1	663	22	ABG30377	Novel human diagno	C 483	8	1.1	28	22	ABBS32896	Peptide #402 enco
411	8	1.1	667	21	AAW17813	Arabidopsis thalia	C 484	8	1.1	28	22	ABBS18376	Protein #375 enco
412	8	1.1	667	21	AAW49206	Arabidopsis thalia	C 485	8	1.1	28	22	AAW53698	Human brain expres
413	8	1.1	670	21	AAW70229	Human RNA-associat	C 486	8	1.1	28	22	AAW66081	Human bone marrow
414	8	1.1	670	22	AAW78950	Human protein SEQ	C 487	8	1.1	28	22	AAW13951	Peptide #385 enco
415	8	1.1	670	22	AAW92989	Human protein sequ	C 488	8	1.1	28	22	AAW26357	Peptide #394 enco
416	8	1.1	671	21	AAW17812	Arabidopsis thalia	C 489	8	1.1	28	22	AAW01694	Peptide #376 enco
417	8	1.1	671	21	AAW49205	Arabidopsis thalia	C 490	8	1.1	28	23	ABG35730	Human peptide enco
C 418	8	1.1	684	21	AAW10556	Human aspartate pr	491	8	1.1	30	23	ABBS41456	Human ovarian anti
419	8	1.1	684	22	AAW79934	Human protein SEQ	492	8	1.1	31	21	AAW51946	Human secreted pro
C 420	8	1.1	691	21	AAW47285	Arabidopsis thalia	C 493	8	1.1	31	21	AAW05988	Peptide #4670 enco
C 421	8	1.1	707	23	AAO14204	Human transporter	494	8	1.1	33	21	AAW39001	Human secreted pep
C 422	8	1.1	717	23	AAU85408	Human protein NOV6	C 495	8	1.1	33	22	AAW50506	Gene #23 associate
C 423	8	1.1	727	22	ABBS71717	Drosophila melanog	C 496	8	1.1	34	22	ABBS28662	Peptide #1333 enco
C 424	8	1.1	750	21	AAW45696	Arabidopsis thalia	C 497	8	1.1	34	22	ABBS33867	Peptide #1373 enco
425	8	1.1	752	20	AAW97799	Streptococcus pneu	C 498	8	1.1	34	22	ABBS40068	Peptide #7574 enco
426	8	1.1	752	20	AAW74407	C. trachomatis gid	C 499	8	1.1	34	22	ABBS19307	Protein #1306 enco
427	8	1.1	752	22	AAW25535	Novel human enzyme	500	8	1.1	34	22	AAW24561	Human brain expres
428	8	1.1	752	22	AAW84212	Amino acid sequenc	C 501	8	1.1	34	22	AAW54631	Human brain expres
429	8	1.1	752	22	AAW84213	Amino acid sequenc	502	8	1.1	34	22	AAW60824	Human bone marrow
430	8	1.1	752	22	AAW79339	Glycogen phosphory	C 503	8	1.1	34	22	AAW67037	Human bone marrow
431	8	1.1	764	23	AAU02736	Bartley DEAD box pr	504	8	1.1	34	22	AAW73504	Human bone marrow
C 432	8	1.1	768	21	AAW45695	Arabidopsis thalia	C 505	8	1.1	34	22	AAW14898	Peptide #1332 enco
433	8	1.1	783	22	AAW39072	Human polypeptide	506	8	1.1	34	22	AAW19993	Peptide #6427 enco

C 507	7	1.0	34	22	AA027327	Peptide #1364 enco	C 580	7	1.0	60	22	AA061926	Human brain expres
C 508	7	1.0	34	22	AA033700	Peptide #7737 enco	C 581	7	1.0	60	22	AA074728	Human bone marrow
C 509	7	1.0	34	22	AA002622	Peptide #1304 enco	C 582	7	1.0	60	22	AA020394	Peptide #6828 enco
C 510	7	1.0	34	23	ABG36694	Human peptide enco	C 583	7	1.0	60	22	AA034845	Peptide #8882 enco
C 511	7	1.0	34	23	ABG43364	Human peptide enco	C 584	7	1.0	60	22	ABG44532	Human peptide enco
C 512	7	1.0	35	22	AA027369	Novel bone marrow	C 585	7	1.0	61	20	AA029365	Fragment of human
C 513	7	1.0	36	20	AA012142	Human 5' EST seque	C 586	7	1.0	62	22	AB016236	Human nervous syst
C 514	7	1.0	37	22	AB044148	Peptide #11648 enc	C 587	7	1.0	62	22	AA090903	Human immune/haema
C 515	7	1.0	37	22	AB027026	Protein #3025 enco	C 588	7	1.0	63	21	AA059384	Arabidopsis thalia
C 516	7	1.0	37	22	AA077874	Human bone marrow	C 589	7	1.0	63	22	AA007988	Human polypeptide
C 517	7	1.0	37	22	AA087275	Human immune/haema	C 590	7	1.0	64	22	ABG29218	Novel human diagno
C 518	7	1.0	37	22	AA021771	Peptide #8205 enco	C 591	7	1.0	64	22	AA098737	Human cell death p
C 519	7	1.0	37	22	AA08095	Peptide #12132 enc	C 592	7	1.0	65	22	ABG06150	Novel human diagno
C 520	7	1.0	37	23	ABG46898	Human polypeptide	C 593	7	1.0	65	22	AA061454	Metallothionein do
C 521	7	1.0	38	22	AA004430	Human polypeptide	C 594	7	1.0	66	21	AA044141	Arabidopsis thalia
C 522	7	1.0	39	22	AA091327	Human immune/haema	C 595	7	1.0	67	22	AA007676	Human polypeptide
C 523	7	1.0	41	10	AA090948	Peptide 242. Foot	C 596	7	1.0	68	21	AA034761	Human secreted pro
C 524	7	1.0	41	20	AA011382	Human 5' EST seque	C 597	7	1.0	68	22	ABG06151	Novel human diagno
C 525	7	1.0	41	22	AB031260	Peptide #3911 enco	C 598	7	1.0	68	22	AB017440	Human nervous syst
C 526	7	1.0	41	22	AB036468	Peptide #3974 enco	C 599	7	1.0	68	23	ABP01087	Human ORFX protein
C 527	7	1.0	41	22	AB021811	Protein #3810 enco	C 600	7	1.0	69	21	AA016529	Arabidopsis thalia
C 528	7	1.0	41	22	AA057230	Human brain expres	C 601	7	1.0	69	21	AA054908	Arabidopsis thalia
C 529	7	1.0	41	22	AA069632	Human bone marrow	C 602	7	1.0	70	23	ABP34865	Human ORP383 prot
C 530	7	1.0	41	22	AA008507	Human polypeptide	C 603	7	1.0	71	22	AA008474	Human polypeptide
C 531	7	1.0	41	22	AA017448	Peptide #3882 enco	C 604	7	1.0	71	22	AA099186	Target molecule hu
C 532	7	1.0	41	22	AA029965	Peptide #4002 enco	C 605	7	1.0	72	20	AA030658	Human secreted pro
C 533	7	1.0	41	22	AA005125	Peptide #3807 enco	C 606	7	1.0	72	21	AA018831	Zea mays protein f
C 534	7	1.0	41	23	ABG39254	Human peptide enco	C 607	7	1.0	72	22	AA054276	Propionibacterium
C 535	7	1.0	42	20	AA011632	Human 5' EST seque	C 608	7	1.0	72	22	ABG00628	Novel human diagno
C 536	7	1.0	43	19	AA075128	Human secreted pro	C 609	7	1.0	72	22	ABG05186	Novel human diagno
C 537	7	1.0	43	22	AB016861	Human nervous syst	C 610	7	1.0	72	22	AA084634	Human immune/haema
C 538	7	1.0	44	20	AA019565	Amino acid sequenc	C 611	7	1.0	72	22	AA023545	Human EST encoded
C 539	7	1.0	45	22	AA092541	Human digestive sy	C 612	7	1.0	73	20	AA019450	Amino acid sequenc
C 540	7	1.0	45	22	AA085857	Human colorectal c	C 613	7	1.0	73	21	AA051998	Arabidopsis thalia
C 541	7	1.0	47	21	AA053805	C-terminal fragmen	C 614	7	1.0	73	22	AB068835	Drosophila melanog
C 542	7	1.0	47	21	AA055227	Arabidopsis thalia	C 615	7	1.0	73	22	AA062145	Propionibacterium
C 543	7	1.0	47	22	AA06103	Human gene 63 enco	C 616	7	1.0	74	21	AA057266	Arabidopsis thalia
C 544	7	1.0	47	23	ABG33925	Human secreted pro	C 617	7	1.0	74	21	AA000565	Human secreted pro
C 545	7	1.0	48	18	AA027289	Human M27-2 secret	C 618	7	1.0	75	21	AA030248	Arabidopsis thalia
C 546	7	1.0	48	18	AA036947	Protein encoded by	C 619	7	1.0	75	22	AA042451	Propionibacterium
C 547	7	1.0	48	19	AA074587	Amino acid sequenc	C 620	7	1.0	75	22	AA089249	Human immune/haema
C 548	7	1.0	48	19	AA048799	C-terminal fragmen	C 621	7	1.0	75	22	AA098736	Human cell death p
C 549	7	1.0	48	19	AA068145	Human AChE splice	C 622	7	1.0	75	23	ABP31124	Human ORP9 protei
C 550	7	1.0	48	21	AA013632	Arabidopsis thalia	C 623	7	1.0	76	22	AA051607	Propionibacterium
C 551	7	1.0	48	21	AA036558	Arabidopsis thalia	C 624	7	1.0	76	22	AA069856	Human prostate CDN
C 552	7	1.0	48	21	AA087126	Human secreted pro	C 625	7	1.0	76	22	AA059933	Human prostate CDN
C 553	7	1.0	48	22	AA082560	Human immune/haema	C 626	7	1.0	76	22	AA010211	Human prostate-spe
C 554	7	1.0	49	22	AA074443	Human colon cancer	C 627	7	1.0	76	22	AA001288	Human prostate-spe
C 555	7	1.0	50	23	ABP31643	Human ORF616 prote	C 628	7	1.0	76	22	AA048733	Mouse liver growth
C 556	7	1.0	50	23	ABP04572	Human ORFX protein	C 629	7	1.0	76	22	AA010192	Human bone marrow
C 557	7	1.0	51	22	AA043078	Propionibacterium	C 630	7	1.0	76	23	AB095316	Human p776P splice
C 558	7	1.0	52	22	ABG27686	Novel human diagno	C 631	7	1.0	76	23	AB095393	Human p776P splice
C 559	7	1.0	52	23	ABP31800	Human ORF773 prote	C 632	7	1.0	77	15	AA062851	Androgen receptor
C 560	7	1.0	53	20	AA014425	Human secreted pro	C 633	7	1.0	77	19	AA048348	Human breast cance
C 561	7	1.0	53	21	AA030376	Novel human secret	C 634	7	1.0	77	22	ABG18992	Novel human diagno
C 562	7	1.0	54	21	AA051558	Arabidopsis thalia	C 635	7	1.0	77	22	AB031623	Peptide #4274 enco
C 563	7	1.0	54	22	AA051558	Propionibacterium	C 636	7	1.0	77	22	AB036840	Peptide #4346 enco
C 564	7	1.0	54	22	AA005053	Human polypeptide	C 637	7	1.0	77	22	AB022163	Protein #4162 enco
C 565	7	1.0	55	20	AA012547	Human 5' EST seque	C 638	7	1.0	77	22	AA057584	Human brain expres
C 566	7	1.0	56	17	AA091275	Elmertia gametocyte	C 639	7	1.0	77	22	AA069988	Human bone marrow
C 567	7	1.0	56	20	AA071912	Partial amino acid	C 640	7	1.0	77	22	AA017819	Peptide #4253 enco
C 568	7	1.0	56	21	AA085728	Novel protein (Cio	C 641	7	1.0	77	22	AA030326	Peptide #4363 enco
C 569	7	1.0	56	21	AA087299	Human signal pepti	C 642	7	1.0	77	22	AA005466	Peptide #4148 enco
C 570	7	1.0	56	22	AA009668	Human polypeptide	C 643	7	1.0	77	23	ABG39617	Human peptide enco
C 571	7	1.0	57	20	AA012935	Amino acid sequenc	C 644	7	1.0	79	21	AA057689	Propionibacterium
C 572	7	1.0	57	22	AA045783	Propionibacterium	C 645	7	1.0	79	22	AA050439	Arabidopsis thalia
C 573	7	1.0	57	22	AA082726	Human immune/haema	C 646	7	1.0	81	21	AA094429	Human Zs1961 prote
C 574	7	1.0	58	22	AA027459	Novel bone marrow	C 647	7	1.0	81	21	AA066619	Human protein havi
C 575	7	1.0	58	22	AA008489	Human polypeptide	C 648	7	1.0	82	21	AA028789	Arabidopsis thalia
C 576	7	1.0	60	19	AA054440	Mouse novel secret	C 649	7	1.0	82	21	AA057688	Arabidopsis thalia
C 577	7	1.0	60	21	AA010247	Murine fecal thymu	C 650	7	1.0	82	22	AA044072	Propionibacterium
C 578	7	1.0	60	22	AB041068	Peptide #8574 enco	C 651	7	1.0	83	22	AB016626	Human nervous syst
C 579	7	1.0	60	22	AB025133	Protein #7132 enco	C 652	7	1.0	83	22	AA009340	Human polypeptide

c 653	7	1.0	83	23	ABP191946	Human ORFp91 prote	726	7	1.0	101	19	AAW53327	Rat Purlipha like
c 654	7	1.0	84	21	AAAG00172	Human secreted pro	727	7	1.0	101	20	AAV36379	Fragment of human
c 655	7	1.0	85	21	AAAG51997	Arabidopsis thalia	728	7	1.0	101	21	AAAG57912	Arabidopsis thalia
c 656	7	1.0	85	22	AAU62629	Propionibacterium	729	7	1.0	101	22	AAO11110	Human polypeptide
c 657	7	1.0	85	22	AAAM24409	Human EST encoded	730	7	1.0	102	20	AAV48446	Human prostate can
c 658	7	1.0	85	22	AAAG76182	Human colon cancer	731	7	1.0	103	22	ABG29245	Novel human diagno
c 659	7	1.0	85	23	AAE23312	Human p456K prote	732	7	1.0	104	21	ABR31003	Pinus radiata tran
c 660	7	1.0	85	23	AAE23318	Human p456K prote	733	7	1.0	104	22	AAAM40702	Human polypeptide
c 661	7	1.0	85	23	ABP29271	Streptococcus poly	734	7	1.0	104	23	ABP05030	Human ORF protein
c 662	7	1.0	86	20	AAV11933	Human 5' EST secre	735	7	1.0	105	21	AAE56137	Human secreted pro
c 663	7	1.0	86	21	AAAB3792	Eucalyptus grandis	736	7	1.0	105	22	AAV70424	Staphylococcus aur
c 664	7	1.0	86	21	AAAG26614	Arabidopsis thalia	737	7	1.0	105	22	AAU36624	Staphylococcus aur
c 665	7	1.0	86	22	ABR28185	Human peptide #836	738	7	1.0	105	22	AAU37167	Staphylococcus aur
c 666	7	1.0	86	22	ABR33360	Peptide #866 encod	739	7	1.0	105	22	AAO05314	Human polypeptide
c 667	7	1.0	86	22	ABR18820	Protein #819 encod	740	7	1.0	106	21	AAAB33907	Human secreted pro
c 668	7	1.0	86	22	AAAM54146	Human brain expres	741	7	1.0	107	22	AAU54125	Propionibacterium
c 669	7	1.0	86	22	AAAM6539	Human bone marrow	742	7	1.0	109	22	ABR59656	Drosophila melanog
c 670	7	1.0	86	22	AAO04805	Human polypeptide	743	7	1.0	109	22	AAU61871	Propionibacterium
c 671	7	1.0	86	22	AAAM14412	Peptide #846 encod	744	7	1.0	109	22	ABG06310	Novel human diagno
c 672	7	1.0	86	22	AAAM26825	Peptide #862 encod	745	7	1.0	109	22	AAE03259	Human gene 6 encod
c 673	7	1.0	86	22	AAAM02139	Peptide #821 encod	746	7	1.0	109	23	ABP38420	Staphylococcus epi
c 674	7	1.0	86	23	ABG36191	Human peptide enco	747	7	1.0	110	22	AAAM83311	Human immune/haema
c 675	7	1.0	87	20	AAV37377	Chlamydia trachoma	748	7	1.0	111	22	AAAB47331	FCR8... Homo sapie
c 676	7	1.0	87	22	ABG29217	Novel human diagno	749	7	1.0	111	23	ABP09196	Human ORF protein
c 677	7	1.0	87	22	AAAM5584	Human immune/haema	750	7	1.0	111	23	AAU83135	Novel secreted pro
c 678	7	1.0	87	23	AAU96975	Human Mch2 protein	751	7	1.0	112	19	AAAM60650	Human C-C chemokin
c 679	7	1.0	88	19	AAU77753	Pyruvate oxidase p	752	7	1.0	112	21	AAAB01453	Primate CTACK prot
c 680	7	1.0	89	16	AAAR75150	Treeshrew (Hiss,Tr	753	7	1.0	112	21	AAV57771	Human CC type chem
c 681	7	1.0	89	22	AAU44512	Propionibacterium	754	7	1.0	112	22	AAO09750	Human polypeptide
c 682	7	1.0	89	22	AAU67315	Propionibacterium	755	7	1.0	113	22	AAU61707	Propionibacterium
c 683	7	1.0	89	22	ABR10828	Human ovarian anti	756	7	1.0	113	23	ABP41890	Human ovarian anti
c 684	7	1.0	89	22	AAAM94499	Human ovariin and/	757	7	1.0	113	23	ABR08792	Human G-substrate
c 685	7	1.0	91	22	AAU414291	Propionibacterium	758	7	1.0	114	21	AAAG54768	Arabidopsis thalia
c 686	7	1.0	91	22	AAU30488	Novel human secret	759	7	1.0	115	22	ABG11703	Novel human diagno
c 687	7	1.0	92	18	AAAM20240	H. pylori secreted	760	7	1.0	118	22	AAU45132	Propionibacterium
c 688	7	1.0	92	21	AAAB42825	Human ORFX ORF2589	761	7	1.0	118	22	AAO07195	Human polypeptide
c 689	7	1.0	92	22	ABG16631	Novel human diagno	762	7	1.0	119	22	AAAM87975	Human immune/haema
c 690	7	1.0	93	19	AAAM53893	Fragment of chimex	763	7	1.0	120	23	AAE15401	Human gonadotropin
c 691	7	1.0	93	22	ABG29222	Novel human diagno	764	7	1.0	121	22	AAAB94940	Human protein sequ
c 692	7	1.0	94	22	ABR17182	Human nervous syst	765	7	1.0	122	23	ABP08628	Human ORFX protein
c 693	7	1.0	95	21	AAAG16528	Arabidopsis thalia	766	7	1.0	122	23	AAU83217	Novel secreted pro
c 694	7	1.0	95	21	AAAG51996	Arabidopsis thalia	767	7	1.0	123	22	AAAM80996	Human haematologic
c 695	7	1.0	95	22	ABR32410	Peptide #5061 enco	768	7	1.0	123	22	AAAM81965	Human haematologic
c 696	7	1.0	95	22	AAAM70797	Human bone marrow	769	7	1.0	124	22	AAU65976	Propionibacterium
c 697	7	1.0	95	22	AAAM06200	Peptide #4882 enco	770	7	1.0	124	22	AAO05830	Human polypeptide
c 698	7	1.0	96	21	AAAG57913	Arabidopsis thalia	771	7	1.0	124	22	AAO05832	Human polypeptide
c 699	7	1.0	96	22	ABR68955	Drosophila melanog	772	7	1.0	124	23	ABP02824	Human ORFX protein
c 700	7	1.0	97	21	AAAG28788	Novel human diagno	773	7	1.0	125	20	AAV13263	Human 5' EST secre
c 701	7	1.0	97	22	ABG29215	Arabidopsis thalia	774	7	1.0	125	21	AAAG26755	Zea mays protein f
c 702	7	1.0	97	22	ABR52761	Escherichia coli p	775	7	1.0	126	22	AAO01854	Human polypeptide
c 703	7	1.0	98	21	AAAB25084	Plant SDF encoded	776	7	1.0	127	21	AAAG24107	Arabidopsis thalia
c 704	7	1.0	98	21	AAAG23146	Arabidopsis thalia	777	7	1.0	128	23	ABP38036	Staphylococcus epi
c 705	7	1.0	98	21	AAAG26613	Arabidopsis thalia	778	7	1.0	129	21	AAAG01659	Human secreted pro
c 706	7	1.0	98	21	AAAG60540	Arabidopsis thalia	779	7	1.0	131	20	AAV37793	Chlamydia trachoma
c 707	7	1.0	98	22	ABR68654	Drosophila melanog	780	7	1.0	131	21	AAV76147	Human secreted pro
c 708	7	1.0	98	22	ABG22290	Novel human diagno	781	7	1.0	131	22	ABG14824	Novel human diagno
c 709	7	1.0	98	22	ABR28261	Human peptide #912	782	7	1.0	132	21	AAAG26612	Arabidopsis thalia
c 710	7	1.0	98	22	ABR33436	Peptide #942 enco	783	7	1.0	133	22	AAO02265	Human polypeptide
c 711	7	1.0	98	22	ABR18895	Protein #894 encod	784	7	1.0	133	23	ABP34915	Human helicase-lik
c 712	7	1.0	98	22	AAAB34221	Human brain expres	785	7	1.0	133	21	AAAG00072	Human secreted pro
c 713	7	1.0	98	22	AAAM66615	Human bone marrow	786	7	1.0	135	22	ABR66419	Drosophila melanog
c 714	7	1.0	98	22	AAAM14488	Peptide #922 encod	787	7	1.0	135	22	AAAG73999	Human colon cancer
c 715	7	1.0	98	22	AAAM26901	Peptide #938 encod	788	7	1.0	136	22	AAO13335	Human polypeptide
c 716	7	1.0	98	22	AAAM38916	Human polypeptide	789	7	1.0	136	22	AAAB36669	Human polypeptide
c 717	7	1.0	98	22	AAAM02215	Peptide #897 encod	790	7	1.0	137	18	AAAM15777	Human secretory pr
c 718	7	1.0	98	23	ABG36267	Human peptide enco	791	7	1.0	137	21	AAAB38331	Human secreted pro
c 719	7	1.0	99	22	AAU34037	Staphylococcus aur	792	7	1.0	137	22	ABR966293	Human testicular a
c 720	7	1.0	99	22	AAU57871	Propionibacterium	793	7	1.0	137	22	ABG081554	Novel human diagno
c 721	7	1.0	99	22	AAAM84606	Human immune/haema	794	7	1.0	137	22	ABG29219	Novel human diagno
c 722	7	1.0	100	18	AAAM20974	H. pylori secreted	795	7	1.0	137	22	AAAM95762	Human reproductive
c 723	7	1.0	100	22	AAU42466	Propionibacterium	796	7	1.0	137	22	AAO094668	Human polypeptide
c 724	7	1.0	100	22	ABG26182	Novel human diagno	797	7	1.0	138	22	AAAB65092	Gene #23 associate
c 725	7	1.0	100	23	ABR89757	Human polypeptide	798	7	1.0	139	21	AAAG57862	Arabidopsis thalia

799	7	1.0	139	23	AB990295	Human polypeptide	872	7	1.0	174	21	AAV74627	Neisseria gonorrhoe
C 800	7	1.0	140	22	AA009801	Human polypeptide	873	7	1.0	174	21	AAV74628	Neisseria meningit
C 801	7	1.0	141	21	AG48936	Arabidopsis thalia	C 874	7	1.0	174	22	AAU23789	Novel human enzyme
C 802	7	1.0	141	21	AG54749	Arabidopsis thalia	C 875	7	1.0	176	19	AAW80254	Amino acid sequenc
C 803	7	1.0	141	22	ABG03186	Novel human thalia	876	7	1.0	176	21	AAW80636	Arabidopsis thalia
C 804	7	1.0	141	22	ABG06227	Novel human thalia	877	7	1.0	176	21	AAW80636	Arabidopsis thalia
C 805	7	1.0	141	23	AAU83201	Novel human thalia	878	7	1.0	176	21	AAW80636	Arabidopsis thalia
C 806	7	1.0	142	21	AAU83201	Novel human thalia	879	7	1.0	176	21	AAW80636	Arabidopsis thalia
C 807	7	1.0	142	21	AAU83201	Novel human thalia	880	7	1.0	177	21	AAW80636	Arabidopsis thalia
C 808	7	1.0	142	22	AAU83201	Novel human thalia	881	7	1.0	177	21	AAW80636	Arabidopsis thalia
C 809	7	1.0	142	22	AAU83201	Novel human thalia	882	7	1.0	177	22	AAW80636	Arabidopsis thalia
C 810	7	1.0	142	22	AAU83201	Novel human thalia	883	7	1.0	177	22	AAW80636	Arabidopsis thalia
C 811	7	1.0	142	22	AAU83201	Novel human thalia	884	7	1.0	177	22	AAW80636	Arabidopsis thalia
C 812	7	1.0	142	23	ABG5429	Human albumin fusi	C 885	7	1.0	179	22	AAU53900	Novel human thalia
C 813	7	1.0	142	23	ABG5429	Human albumin fusi	C 886	7	1.0	180	22	AAU53900	Novel human thalia
C 814	7	1.0	143	21	AG17023	Arabidopsis thalia	C 887	7	1.0	180	22	AAU53900	Novel human thalia
C 815	7	1.0	143	21	AG17023	Arabidopsis thalia	C 888	7	1.0	180	22	AAU53900	Novel human thalia
C 816	7	1.0	143	22	AAU53900	Human ngPCR55. Ho	C 889	7	1.0	180	22	AAU53900	Novel human thalia
C 817	7	1.0	144	15	AAU53900	Human ngPCR55. Ho	C 890	7	1.0	180	22	AAU53900	Novel human thalia
C 818	7	1.0	144	15	AAU53900	Human ngPCR55. Ho	C 891	7	1.0	180	22	AAU53900	Novel human thalia
C 819	7	1.0	144	21	AAU53900	Human ngPCR55. Ho	C 892	7	1.0	180	22	AAU53900	Novel human thalia
C 820	7	1.0	144	22	AAU53900	Human ngPCR55. Ho	C 893	7	1.0	180	22	AAU53900	Novel human thalia
C 821	7	1.0	145	22	AAU53900	Human ngPCR55. Ho	C 894	7	1.0	181	22	AAU53900	Novel human thalia
C 822	7	1.0	146	22	AAU53900	Human ngPCR55. Ho	C 895	7	1.0	181	22	AAU53900	Novel human thalia
C 823	7	1.0	147	20	AAU53900	Human ngPCR55. Ho	C 896	7	1.0	181	22	AAU53900	Novel human thalia
C 824	7	1.0	147	22	AAU53900	Human ngPCR55. Ho	C 897	7	1.0	182	22	AAU53900	Novel human thalia
C 825	7	1.0	147	22	AAU53900	Human ngPCR55. Ho	C 898	7	1.0	183	22	AAU53900	Novel human thalia
C 826	7	1.0	147	22	AAU53900	Human ngPCR55. Ho	C 899	7	1.0	183	22	AAU53900	Novel human thalia
C 827	7	1.0	149	23	ABG0343	Human polypeptide	C 900	7	1.0	183	22	AAU53900	Novel human thalia
C 828	7	1.0	150	23	ABG0343	Human polypeptide	C 901	7	1.0	183	22	AAU53900	Novel human thalia
C 829	7	1.0	150	23	ABG0343	Human polypeptide	C 902	7	1.0	183	22	AAU53900	Novel human thalia
C 830	7	1.0	150	23	ABG0343	Human polypeptide	C 903	7	1.0	183	22	AAU53900	Novel human thalia
C 831	7	1.0	151	21	AAU53900	Human polypeptide	C 904	7	1.0	184	21	AAU53900	Novel human thalia
C 832	7	1.0	152	21	AAU53900	Human polypeptide	C 905	7	1.0	184	21	AAU53900	Novel human thalia
C 833	7	1.0	152	22	AAU53900	Human polypeptide	C 906	7	1.0	185	20	AAU53900	Novel human thalia
C 834	7	1.0	152	22	AAU53900	Human polypeptide	C 907	7	1.0	185	20	AAU53900	Novel human thalia
C 835	7	1.0	153	22	AAU53900	Human polypeptide	C 908	7	1.0	188	22	AAU53900	Novel human thalia
C 836	7	1.0	154	22	AAU53900	Human polypeptide	C 909	7	1.0	190	20	AAU53900	Novel human thalia
C 837	7	1.0	154	22	AAU53900	Human polypeptide	C 910	7	1.0	191	22	AAU53900	Novel human thalia
C 838	7	1.0	154	22	AAU53900	Human polypeptide	C 911	7	1.0	191	22	AAU53900	Novel human thalia
C 839	7	1.0	155	22	AAU53900	Human polypeptide	C 912	7	1.0	192	16	AAU53900	Novel human thalia
C 840	7	1.0	155	22	AAU53900	Human polypeptide	C 913	7	1.0	192	16	AAU53900	Novel human thalia
C 841	7	1.0	155	22	AAU53900	Human polypeptide	C 914	7	1.0	192	16	AAU53900	Novel human thalia
C 842	7	1.0	157	21	AAU53900	Human polypeptide	C 915	7	1.0	192	17	AAU53900	Novel human thalia
C 843	7	1.0	157	21	AAU53900	Human polypeptide	C 916	7	1.0	192	21	AAU53900	Novel human thalia
C 844	7	1.0	157	21	AAU53900	Human polypeptide	C 917	7	1.0	192	21	AAU53900	Novel human thalia
C 845	7	1.0	158	21	AAU53900	Human polypeptide	C 918	7	1.0	193	22	AAU53900	Novel human thalia
C 846	7	1.0	158	21	AAU53900	Human polypeptide	C 919	7	1.0	193	22	AAU53900	Novel human thalia
C 847	7	1.0	158	22	AAU53900	Human polypeptide	C 920	7	1.0	193	23	AAU53900	Novel human thalia
C 848	7	1.0	159	20	AAU53900	Human polypeptide	C 921	7	1.0	194	13	AAU53900	Novel human thalia
C 849	7	1.0	159	21	AAU53900	Human polypeptide	C 922	7	1.0	194	13	AAU53900	Novel human thalia
C 850	7	1.0	159	21	AAU53900	Human polypeptide	C 923	7	1.0	196	21	AAU53900	Novel human thalia
C 851	7	1.0	159	22	AAU53900	Human polypeptide	C 924	7	1.0	196	21	AAU53900	Novel human thalia
C 852	7	1.0	160	21	AAU53900	Human polypeptide	C 925	7	1.0	196	21	AAU53900	Novel human thalia
C 853	7	1.0	160	21	AAU53900	Human polypeptide	C 926	7	1.0	197	21	AAU53900	Novel human thalia
C 854	7	1.0	160	22	AAU53900	Human polypeptide	C 927	7	1.0	197	21	AAU53900	Novel human thalia
C 855	7	1.0	160	22	AAU53900	Human polypeptide	C 928	7	1.0	197	21	AAU53900	Novel human thalia
C 856	7	1.0	161	22	AAU53900	Human polypeptide	C 929	7	1.0	197	22	AAU53900	Novel human thalia
C 857	7	1.0	161	23	AAU53900	Human polypeptide	C 930	7	1.0	197	22	AAU53900	Novel human thalia
C 858	7	1.0	163	23	AAU53900	Human polypeptide	C 931	7	1.0	198	21	AAU53900	Novel human thalia
C 859	7	1.0	164	21	AAU53900	Human polypeptide	C 932	7	1.0	198	21	AAU53900	Novel human thalia
C 860	7	1.0	166	22	AAU53900	Human polypeptide	C 933	7	1.0	199	16	AAU53900	Novel human thalia
C 861	7	1.0	166	22	AAU53900	Human polypeptide	C 934	7	1.0	200	22	AAU53900	Novel human thalia
C 862	7	1.0	167	21	AAU53900	Human polypeptide	C 935	7	1.0	201	22	AAU53900	Novel human thalia
C 863	7	1.0	167	21	AAU53900	Human polypeptide	C 936	7	1.0	202	20	AAU53900	Novel human thalia
C 864	7	1.0	168	21	AAU53900	Human polypeptide	C 937	7	1.0	202	20	AAU53900	Novel human thalia
C 865	7	1.0	170	16	AAU53900	Human polypeptide	C 938	7	1.0	202	21	AAU53900	Novel human thalia
C 866	7	1.0	170	21	AAU53900	Human polypeptide	C 939	7	1.0	202	21	AAU53900	Novel human thalia
C 867	7	1.0	170	22	AAU53900	Human polypeptide	C 940	7	1.0	203	16	AAU53900	Novel human thalia
C 868	7	1.0	171	16	AAU53900	Human polypeptide	C 941	7	1.0	203	16	AAU53900	Novel human thalia
C 869	7	1.0	173	21	AAU53900	Human polypeptide	C 942	7	1.0	203	21	AAU53900	Novel human thalia
C 870	7	1.0	173	21	AAU53900	Human polypeptide	C 943	7	1.0	204	22	AAU53900	Novel human thalia
C 871	7	1.0	173	21	AAU53900	Human polypeptide	C 944	7	1.0	205	22	AAU53900	Novel human thalia

945	7	1.0	205	23	AAE14609	Human microtubule
946	7	1.0	207	21	AAE424073	Zea mays protein f
947	7	1.0	207	21	AAE42186	Arabidopsis thalia
948	7	1.0	207	22	AAU49163	Propionibacterium
949	7	1.0	210	21	AAE11097	Arabidopsis thalia
950	7	1.0	210	22	ABG28848	Novel human diago
951	7	1.0	212	20	AAU01198	Polypeptide fragme
952	7	1.0	213	20	ABE67036	Drosophila melanog
953	7	1.0	214	20	AAU06303	Mouse pancreatic-d
954	7	1.0	214	22	ABG12832	Novel human diago
955	7	1.0	216	17	AAE88670	Fusion protein 52/
956	7	1.0	216	21	AAE51946	Arabidopsis thalia
957	7	1.0	216	22	ABE6796	Drosophila melanog
958	7	1.0	217	21	AAE17021	UL26 protease dele
959	7	1.0	218	13	AAE28645	Drosophila melanog
960	7	1.0	218	22	ABE66933	Human G protein-co
961	7	1.0	218	22	AAU19191	Neisseria meningit
962	7	1.0	219	21	AAE75383	Neisseria meningit
963	7	1.0	219	21	AAE75384	Rat 3395347/77-324
964	7	1.0	219	23	AAE15935	Arabidopsis thalia
965	7	1.0	220	21	AAE08365	Propionibacterium
966	7	1.0	222	22	AAU61742	Novel human diago
967	7	1.0	222	22	ABG29350	Human olfactory re
968	7	1.0	222	22	AAE72802	Arabidopsis cell c
969	7	1.0	222	23	AAU72539	Listeria monocytog
970	7	1.0	223	23	AAE29539	Human 51-3 protein
971	7	1.0	223	23	ABE48116	Human zinc finger
972	7	1.0	224	20	AAU09535	Human 4502341_JBC
973	7	1.0	224	22	AAE78055	Drosophila melanog
974	7	1.0	226	23	AAE15934	Novel human diago
975	7	1.0	227	22	ABE69369	Rice invertase inh
976	7	1.0	228	22	ABG18908	Mouse EKR2 protein
977	7	1.0	228	22	AAU05775	Arabidopsis thalia
978	7	1.0	228	22	AAE05346	Novel human diago
979	7	1.0	229	21	AAE08364	Arabidopsis thalia
980	7	1.0	231	22	ABG26553	Novel human diago
981	7	1.0	232	21	AAE21411	Arabidopsis thalia
982	7	1.0	232	21	AAE60408	Product of Q regio
983	7	1.0	233	8	AAE71668	Novel human diago
984	7	1.0	236	22	ABG05159	Streptococcus poly
985	7	1.0	236	23	ABE26618	Arabidopsis thalia
986	7	1.0	237	21	AAE6735	Drosophila melanog
987	7	1.0	237	22	ABE58541	Drosophila melanog
988	7	1.0	237	22	ABE67675	UL26 protease dele
989	7	1.0	238	13	AAE28639	Arabidopsis thalia
990	7	1.0	239	21	AAE07732	Human albumin fusi
991	7	1.0	240	22	AAE90252	Human immune/haema
992	7	1.0	240	22	AAE03207	Human gene 6 encod
993	7	1.0	240	22	AAE03225	Human gene 6 encod
994	7	1.0	240	22	AAE88389	Human membrane or
995	7	1.0	240	23	ABE64408	Human albumin fusi
996	7	1.0	240	23	ABE64409	Human albumin fusi
997	7	1.0	240	23	ABE9440	Listeria monocytog
998	7	1.0	241	23	ABE70942	Drosophila melanog
999	7	1.0	242	21	AAE08363	Arabidopsis thalia
1000	7	1.0	242	21	AAE08674	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
AAE02417  
ID AAE02417 standard; Protein; 724 AA.

XX AAE02417;

DT 10-AUG-2001 (first entry)

DE Human vasa protein.

XX Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;  
KM ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;  
KM colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;

KM	skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;
KM	medulloblastoma; chorioepithelioma; squamous cell carcinoma; leukaemia;
KM	acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;
KM	osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;
KM	leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;
KM	fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;
KM	teratoma; mediastinal; intracranial.
OS	Homo sapiens.
XX	
PH	Key
FT	Region
FT	Region
FT	Region
FT	Region
FT	Region
XX	
XX	MO200136445-A1.
PN	25-MAY-2001.
XX	16-NOV-2000; 2000WO-US31485.
XX	18-NOV-1999; 99US-0166394.
PR	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.
PA	Gastrillon DH;
XX	WPI; 2001-355606/37.
XX	N-PSDB; AAD06354.
DR	Novel vasa polynucleotides useful in the diagnosis or treatment of
PT	conditions characterized by aberrant expression and/or presence of
PT	mutant forms of vasa polynucleotides or polypeptides -
XX	
XX	Claim 18; Page 53-54; 66pp; English.
PS	
XX	The present sequence is human vasa protein that has germ cell specific
CC	expression and is believed to play a determinative role in gonad
CC	development. Germ cells are specialised to produce haploid gametes in
CC	multicellular organisms. Vasa is useful in the diagnosis or treatment
CC	of conditions characterised by its aberrant expression and/or the
CC	presence of its mutant forms. The conditions include cancers such as
CC	biliary tract, brain, breast, colon, ovarian, pancreas, prostate,
CC	colorectal, oral, liver, lung, skin, basocellular, testis, renal,
CC	thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,
CC	melanomas, glioblastomas, neuroblastomas, medulloblastomas,
CC	choriocarcinoma, squamous cell carcinoma, haematological neoplasms,
CC	acute lymphocytic and myelogenous leukaemia, multiple myeloma,
CC	Acquired immune deficiency syndrome (AIDS) associated leukaemias,
CC	intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas
CC	such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,
CC	Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour
CC	(eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour
CC	of an extragonadal tissue (eg. mediastinal or an intracranial tumour).
XX	
SQ	Sequence 724 AA;

## Alignment Scores:

Pred. No.:	0	Length:	724
Score:	724.00	Matches:	724
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-714-865-15 (1-2172) x AAE02417 (1-724)

QY 1 ATGGGGGATGACATTGGGAGGAGCAAAATCAACCTTCATATGTTCTTCATATGTTCCATA 60  
Db 1 MetGlyAspGluAspTrpGluAlaGluIleAsnProHisMetSerSerYrValProIle 20



QY 61 TTGAGAGGATAGCTATTTCTGGAGAAATGGAGACAAATTTTAAACAGACTCCAGCTTCA 120  
 Db 21 PheGluysAspArgTyrSerGlyGluAsnGlyAspAsnPhenAlaGThrProAlaSer 40  
 QY 121 TCATCGAATATGATGATGACCTTCTCGAAGAGATCATATTTCATGAAAATGGATTTGCC 180  
 Db 41 SerSerGluMetAspArgGlyProSerArgRgsAspHisPheMetCysSerGlyPheAla 60  
 QY 181 TCTGGCGGCAATTTTGGAAAACAGAGATGCTGTGAGTGTAAATAAGCAGATTAATATCATCC 240  
 Db 61 SerGlyArgAsnPhenGlyAsnArgAspAlaGlyGlyCysAsnLysArgAspAsnThrSer 80  
 QY 241 ACAATGGGTGTTTGGAGTTGGAAAAGATTTTGGAAAACAGAGGTTTTCACACAGCAGG 300  
 Db 81 ThrMetGlyGlyPheGlyValGlyLysSerPheGlyLeuAsnArgGlyPheSerAsnSerArg 100  
 QY 301 TTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
 Db 101 PheGluAspGlyAspSerSerGlyPheTrpArgGlyLysSerAsnAspCysGluAspAsn 120  
 QY 361 CCAACAGGAGACAGAGGTTTCCAGAGAGGCGCTATGAGATGAGATTAATTTTCGAA 420  
 Db 121 ProThrArgAsnArgGlyPheSerLysArgGlyLysTrpArgAspGlyAsnAsnSerGlu 140  
 QY 421 GCTTCAGGCGCATACAGAAAGGTGGAAGAGTGTAGTTCCGAGGTTCCGTGAGAGATT 480  
 Db 141 AlaSerGlyProTyrArgArgGlyGlyArgGlySerPheArgGlyCysArgGlyGlyPhe 160  
 QY 481 GGCTTCAGGAAGTCCAAATTAATGACTTAAGCCACAGACAAATGATGACCGCATGTTGGC 540  
 Db 161 GlyLeuGlySerProAsnAsnAspLeuAspProAspGlyCysMetGlnArgThrGlyGly 180  
 QY 541 CTTTGTGCTCTGAAGACAGATTAATGAGTGGACAGTAAAGTGTACTCTCCAGAC 600  
 Db 181 LeuPheGlySerArgArgProValLeuSerGlyThrGlyAsnGlyAspThrSerGlnSer 200  
 QY 601 AGAAGTGGCAGTGAAGTGAACGAGTGTGTTACAAAGTTTAAATGAAGATTAATACA 660  
 Db 201 ArgSerGlySerGlySerGluArgGlyGlyTyrLysGlyLeuAsnGluGluValIleThr 220  
 QY 661 GGCTTCGAAAAGAAATTTGGAAGTCAAGACAGAGAGAGAGAAAGTATGATCTCAA 720  
 Db 221 GlySerGlyLysAsnSerTrpLysSerGlyLysGlyGlyGlyGlySerSerAspThrGln 240  
 QY 721 GGACCAAAAGTGAAGTCAATACCCCTCCACTGAGATGAGAGTCCATCTTTCGA 780  
 Db 241 GlyProLysValIleThrTyrIleProProProProGluAspGlyAspSerIlePheAla 260  
 QY 781 CATATCAGACAGGATTAATCTGCAAAATACGACACTATTTCTTGGAAAGTGTGGA 840  
 Db 261 HisTyrGlnThrGlyIleAsnPhenAspLysTyrAspThrIleLeuValGlyValSerGly 280  
 QY 841 CATGATCACCACACAGCAATTTGACTTTTGAAGACCTAATCTGTGCACAGCTGAAT 900  
 Db 281 HisAspAlaProProAlaIleLeuThrPheGlyGluLysAsnLeuGlyGlnThrLeuAsn 300  
 QY 901 AACCAACTTGTCTAAAGTGTGTTATACCTAAGCTTACTCTGTGCAAAAATACAGTATCT 960  
 Db 301 AsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIlePro 320  
 QY 961 ATCATACTTCAGAGCAGAGATTTGATGAGTGTGCTGCAACAGAGGTCTGGGAAACAGCGC 1020  
 Db 321 IleIleLeuAlaGlyArgAspLeuMetAlaCysAlaGlnThrGlyLysSerGlyLysThrAla 340  
 QY 1021 GCTTTTCTCTCAATTTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 Db 341 AlaPheLeuLeuProIleLeuAlaHisMetMetHisAspGlyIleThrAlaSerArgPhe 360  
 QY 1081 AAAGAGTTGACAGAAACAGAGTGTATATTTCTACACACACTGAGAAATTTGTAACAG 1140  
 Db 361 LysGluLeuGlnGluProGluCysIleIleValAlaProThrArgGluLeuValAsnGln 380

QY 1141 ATTTATTTGGAACCCAGAAAATTTCTTTTGGACCTTGCTGAAGACTGTGTTATATAT 1200  
 Db 381 IleTyrLeuGlnLysAlaArgLysPheSerPheGlyThrCysValArgAlaValIleTyr 400  
 QY 1201 GGGGGAACCCAGCTGGGACATTTCAATTCGAAATATGTCACAGCGTGAATATATATGCT 1260  
 Db 401 GlyGlyThrGlnLeuGlyHisSerIleArgGlnIleValGlnGlnCysAsnIleLeuCys 420  
 QY 1261 GCTACTCTCGGAAGACTGATGATATCATAGCAAGAAAAGAAAGATTTGCTCAACAGATC 1320  
 Db 421 AlaThrProGlyArgLeuMetAspIleIleGlyLysGlyLysIleGlyLeuLysGlnIle 440  
 QY 1321 AAATACTTACTTTTGGATGAAGCTGATCGCATGTTGATATGAGTGTGTTGCTCAGAAATG 1380  
 Db 441 LysTyrLeuValLeuAspGlyLysAlaAspArgMetLeuAspMetGlyPheGlyProGluMet 460  
 QY 1381 AAGAAGTAAATTTCTTGCCCAAGAAAGCCATCAAAAGAAACAGGCGCAACCTTATGCTTC 1440  
 Db 461 LysLeuLeuIleSerCysProGlyMetProSerLysGlnGlnArgGlnThrLeuMetPhe 480  
 QY 1441 AGTGAACCTTTTCCAGAGAAATTTCAAGAGTGTGCTGACAGTGTTTTAAAGTCAATAT 1500  
 Db 481 SerAlaThrPheProGluGlnIleGlnArgLeuAlaGluPheLeuLysSerAsnTyr 500  
 QY 1501 CTGTTTGTGCTGTTGACACAGTGGGTGAGCATGTAGAGATGTTACAGACCGTTCTC 1560  
 Db 501 LeuPheValAlaValGlyGlnValGlyGlyAlaCysArgAspValGlnGlnThrValLeu 520  
 QY 1561 CAAGTTGGCCAGTTCTCAAAAAGAAAGCTGTTGAAATTTCTGCAACATAGGGGAT 1620  
 Db 521 GlnValGlyGlnPheSerLysArgGlyLysLeuValGlnIleLeuArgAsnIleGlyAsp 540  
 QY 1621 GAAAGAACTATGGCTTTGTTGAACATAAGAAAAGCAGATTTTCTGCACTTTTCTT 1680  
 Db 541 GluArgThrMetValPheValGluThrLysLysLysAlaAspPheThrAlaThrPheLeu 560  
 QY 1681 TGTCAAGAAAATATCAACTACAGATATCCATGATGATCGGGAACAGAGACCGGAG 1740  
 Db 561 CysGlnGlnLysIleSerThrThrSerIleHisGlyAspArgGlyGlnArgGluValGlu 580  
 QY 1741 CAAGCTTTGAGAGATTTTGGCTTTGGAAGTGGCCAGTTCTTGTGCTACTTCAATAGCT 1800  
 Db 581 GlnAlaLeuGlyAspPheArgPheGlyLysCysProValLeuValAlaThrSerValAla 600  
 QY 1801 GCCAGAGGCTGATATTTGAATAATGTCACAAATGATATCAATTTTATCTCTCTGACC 1860  
 Db 601 AlaArgGlyLeuAspIleGlyLysAsnValGlnHisValIleAsnPhenAspLeuProSerThr 620  
 QY 1861 ATTGATGAATATGTCATCGAATTTGGGCGTACTGTCGTTGTGGAATACTGACAGACA 1920  
 Db 621 IleAspGlyLysValHisArgIleGlyArgThrGlyArgCysGlyAsnThrGlyArgAla 640  
 QY 1921 ATTTCTTTTGTATCTTGATCGAATACCAATTAGCACAGCCTCTAGTAAAGTATG 1980  
 Db 641 IleSerPhePheAspLeuGlnSerAspAsnHisLeuAlaGlnProLeuValLysValLeu 660  
 QY 1981 ACAGATGCTCAACAGAGATGTCCTGCATGTTGGAAGAAATTCCTTATCATATCAT 2040  
 Db 661 ThrAspAlaGlnGlnAspValProAlaTrpLeuGlnGlnIleAlaPheSerThrTyrIle 680  
 QY 2041 CTTGCTTCAGTGTGATCAAGAGAAAGCTTTTGCATCACTTGAATACCAAGAGGCG 2100  
 Db 681 ProGlyPheSerGlySerThrArgGlyAsnValPheAlaSerValAspThrArgLysGly 700  
 QY 2101 AAGAGCATTTGAACAACGCTGGGTTTCTTTCACAGAGCTCCAAATCCAGTATGAT 2160  
 Db 701 LysSerThrLeuAsnThrAlaGlyPheSerSerSerArgAlaProAsnProValAspAsp 720  
 QY 2161 GACTCATGGGAT 2172  
 Db 721 GluSerTrpAsp 724

RESULT 2

PR	08-SEP-2000	2000US-0233080
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0232397
PR	14-SEP-2000	2000US-0232398
PR	14-SEP-2000	2000US-0232399
PR	14-SEP-2000	2000US-0232400
PR	14-SEP-2000	2000US-0232401
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	25-SEP-2000	2000US-0234977
PR	25-SEP-2000	2000US-0234998
PR	26-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235634
PR	27-SEP-2000	2000US-0235635
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	02-OCT-2000	2000US-0236702
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239355
PR	13-OCT-2000	2000US-0239357
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241186
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241897
PR	20-OCT-2000	2000US-0241809
PR	01-NOV-2000	2000US-0244617
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246528
PR	11-NOV-2000	2000US-0249607
PR	11-NOV-2000	2000US-0249608
PR	11-NOV-2000	2000US-0249609
PR	11-NOV-2000	2000US-0249610
PR	11-NOV-2000	2000US-0249611
PR	11-NOV-2000	2000US-0249612
PR	11-NOV-2000	2000US-0249613
PR	11-NOV-2000	2000US-0249614
PR	11-NOV-2000	2000US-0249615
PR	11-NOV-2000	2000US-0249616
PR	11-NOV-2000	2000US-0249617
PR	11-NOV-2000	2000US-0249618
PR	11-NOV-2000	2000US-0249619
PR	11-NOV-2000	2000US-0249620
PR	11-NOV-2000	2000US-0249621
PR	11-NOV-2000	2000US-0249622
PR	11-NOV-2000	2000US-0249623
PR	11-NOV-2000	2000US-0249624
PR	11-NOV-2000	2000US-0249625
PR	11-NOV-2000	2000US-0249626
PR	11-NOV-2000	2000US-0249627
PR	11-NOV-2000	2000US-0249628
PR	11-NOV-2000	2000US-0249629
PR	11-NOV-2000	2000US-0249630
PR	11-NOV-2000	2000US-0249631
PR	11-NOV-2000	2000US-0249632
PR	11-NOV-2000	2000US-0249633
PR	11-NOV-2000	2000US-0249634
PR	11-NOV-2000	2000US-0249635
PR	11-NOV-2000	2000US-0249636
PR	11-NOV-2000	2000US-0249637
PR	11-NOV-2000	2000US-0249638
PR	11-NOV-2000	2000US-0249639
PR	11-NOV-2000	2000US-0249640
PR	11-NOV-2000	2000US-0249641
PR	11-NOV-2000	2000US-0249642
PR	11-NOV-2000	2000US-0249643
PR	11-NOV-2000	2000US-0249644
PR	11-NOV-2000	2000US-0249645
PR	11-NOV-2000	2000US-0249646
PR	11-NOV-2000	2000US-0249647
PR	11-NOV-2000	2000US-0249648
PR	11-NOV-2000	2000US-0249649
PR	11-NOV-2000	2000US-0249650
PR	11-NOV-2000	2000US-0249651
PR	11-NOV-2000	2000US-0249652
PR	11-NOV-2000	2000US-0249653
PR	11-NOV-2000	2000US-0249654
PR	11-NOV-2000	2000US-0249655
PR	11-NOV-2000	2000US-02

PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465566/50.  
DR N-PSDB; AAS40979.

XX Novel polypeptides and polynucleotides useful for diagnosing,  
PT preventing, treating neural, immune system, muscular, reproductive,  
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
PT diseases

PS Claim 11; SEQ ID NO 1105; 1180bp; English.

XX The present invention relates to the isolation of novel human enzyme  
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
CC isomerases or ligases. The sequences of the invention are useful in the  
CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
CC disorders including hyperproliferative disorders (e.g. cancer),  
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
CC blood-related disorders (e.g. haemophilia), reproductive disorders  
CC (e.g. infertility) and infectious disorders (e.g. influenza). The  
CC polynucleotides of the invention can also be used in gene therapy.  
CC AAU22915-AAU23814 represent the novel human enzyme polypeptides of the  
CC invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 135 AA;

Alignment Scores:

Pred. No.: 9.91e-111 Length: 135  
Score: 123.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 16.99% Indels: 0  
DB: 22 Gaps: 0

US-09-714-865-15 (1-2172) x AAU23109 (1-135)

QY 1768 AAGTCCAGCTTCTTGTGCTACTTCACTAGTCCAGAGGGCTGATTAATGAATGTG 1827

Db 1 LysCysProValLeuValAlaThrSerValAlaAlaArgLysLeuAspIleGluAsnVal 20

QY 1828 CAACATGTTATCAATTTTGTATCTTCTTCAACATGATGAATATGTTTCATCGAATTGGG 1887

Db 21 GlnHisValIleAsnPheAspLeuProSerThrIleAspGluTyrValHisArgIleGly 40

QY 1888 CGTACTGCTGCTTGGGAATACAGAGACATTTCTTTTGTATCTTGAATCGGAT 1947

Db 41 ArgThrGlyArgCysGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGluSerAsp 60

QY 1948 AACCATTTAGCACAGCCTCTAGTAAAGATTTGACAGATGCTCAACAGATGTTCTCGCA 2007

Db 61 AsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnIleAsnValProAla 80

QY 2008 TGGTTGGAAATAATTCGCTTTACTATACATTCCTCGCTTCGTGTTACTACAAAGCA 2067

Db 81 TrpLeuGlnIleValLeuAlaPheSerThrTyrIleProGlyPheSerGlySerThrArgGly 100

QY 2068 AACGTGTTTCATCAGTTGATACCAAGAGGCAAGACACTTTGAACACAGCTGGGTTT 2127

Db 101 AsnValPheAlaSerValAspThrArgLysGlyLysSerThrLeuAsnThrAlaGlyPhe 120

QY 2128 TCTTCTTCA 2136

Db 121 SerSerSer 123

RESULT 3

ABB95898 ABB95898 standard; Protein; 106 AA.

XX ABB95898;

DT 21-JUN-2002 (first entry)

XX Human testicular antigen SEQ ID NO: 1282.

XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;

XX reproductive system disorder; urinary system disorder; gene therapy;

XX cardiovascular disorder; respiratory disorder; neurological disorder;

XX gastrointestinal disease; infection; cytostatic.

XX Homo sapiens.

XX WO200155317-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01329.

XX 31-JUN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUL-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224516.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 16-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226686.

XX 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0237009.  
PR 30-AUG-2000; 2000US-0238924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235835.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239393.  
PR 13-OCT-2000; 2000US-0239397.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
DR WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating testicular cancer  
XX  
XX Claim 11; SEQ ID NO 1282; 766pp; English.  
PS  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a protein of the  
CC invention.  
XX  
SQ Sequence 106 AA;  
  
Alignment Scores: 1.09e-72 Length: 106  
Pred. No.: 84.00 Matches: 84  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 11.60% Indels: 0  
Query Match: 22 Gaps: 0  
DB:  
  
US-09-714-865-15 (1-2172) x ABB95698 (1-106)  
QY 1885 GGGGCTACTGTCGTTGTGGGAATCTGCAGAGCAATTCCTTTTGAATCTGAATCG 1944  
|||  
Db 11 G1yArghrIglYArGcYsG1yAerThnG1yArghAlIeserRherneApIeugIser 30  
|||  
QY 1945 GATACCATTTAGACAGCCTTAGTAAAGATTTGACAGATGCTCAACAGATGTTCT 2004  
|||  
Db 31 AspAenHIsLeuAlaGInPoleuValIyValIeuthrAspIaGInAAspValPro 50  
|||  
QY 2005 GCATGTTGAGAAATTCCTTTAGTACATCACTTCTGCTTCAGTGGTACAGACAGA 2064  
|||  
Db 51 AlAtIpLeuGInuIleAlaPheSerThrTyrlIeProGlyPheSerGlySerThrArg 70  
|||

Qy 2065 GGAACGCTTTGATAGTTGATACGAAAGGCGAAGACCTTTGAAACACAGCTGGG 2124  
Db 71 GlyAsnValPheAsValAspThrArgLysSerThrLeuAsnThrAlaGly 90  
Qy 2125 TTTTCTTCTTCA 2136  
Db 91 PheSerSerSer 94  
RESULT 4  
AAU23664  
ID AAU23664 standard; Protein; 106 AA.  
AAU23664;  
AC  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human enzyme polypeptide #750.  
XX  
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;  
KW nephrotoxic; anticoagulant.  
XX  
XX  
OS Homo sapiens.  
XX  
PN WQ200155301-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01239.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234598.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.



```

PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241836.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251868.
PR 05-DEC-2000; 2000US-0251869.
PR 05-DEC-2000; 2000US-0251888.
PR 05-DEC-2000; 2000US-0251889.
PR 05-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-46570/50.
XX DR N-PSDB; AAL01169.
XX PT Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition
XX
XX PS Claim 11; SEQ ID NO 3857; 1297bp + Sequence Listing; English.
XX
XX CC The present invention provides the protein and coding sequences of a
XX CC number of human reproductive system related antigens. These can be used
XX CC in the prevention and treatment of reproductive system disorders,
XX CC including cancer. The present sequence is a protein of the invention.
XX
SQ Sequence 106 AA:
SQ
Alignment Scores:
Pred. No.: 1,09e-72 Length: 106
Score: 84.00 Matches: 84
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.60% Indels: 0
DB: 22 Gaps: 0

US-09-714-865-15 (1-2172) x AAM95199 (1-106)
QY 1885 GGGCGTACTGCTGCTGTGGGAATCTGGCAGCAATTCCTTTTGTGATCTTGAATCG 1944
Db 11 GlyArgThrGlyArgGlyAsnThrGlyArgAlaIleSerPhePheAspLeuGlnSer 30
QY 1945 GATACCATTTTACACAGCCTCTAGTAAAGTATTACACGATGCTCAACAGATGTTCT 2004
Db 31 AspAsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnAspValPro 50
QY 2005 GCATGTTGGAGAAATTCCTTTAGTACATACATTCCTGCTTCACTAGTACACAGA 2064
Db 51 AlaTrpLeuGlnIuIleAlaPheSerThrTyIleProGlyPheSerGlySerThrArg 70
QY 2065 GGAACGTTGTTTCATCATGTTGATACACAGGAGGAGGACCTTTGAACACAGCTGGG 2124
Db 71 GlyAsnValPheAlaSerValaPheThrArgGlyGlySerThrLeuAsnThrAlaGly 90

```

Qy 2125 TTTTCTCTCA 2136  
 Db 91 PheserSerSer 94  
 RESULT 6  
 AAE02419  
 ID AAE02419 standard; Protein: 713 AA.  
 AC AAE02419;  
 XX  
 DT 10-AUG-2001 (first entry)  
 XX  
 DE Rat vasa protein.  
 XX  
 KM Rat; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;  
 KM pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;  
 KM thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;  
 KM melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;  
 KM choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;  
 KM myelogenous; multiple myeloma; Paget's disease; osteosarcoma;  
 KM Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
 KM leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
 KM fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
 KM teratoma; mediastinal; intracranial.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN WO200136445-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 16-NOV-2000; 2000WO-US31485.  
 XX  
 PR 18-NOV-1999; 99US-0166394.  
 XX  
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.  
 XX  
 PI Castriillon DH;  
 XX  
 XX WPI; 2001-355606/37.  
 XX  
 DR Novel vasa polynucleotides useful in the diagnosis or treatment of  
 PT conditions characterized by aberrant expression and/or presence of  
 PT mutant forms of vasa polynucleotides or polypeptides -  
 PT  
 Claim 4; Page 56-58; 66pp; English.  
 XX  
 PS The present sequence is rat vasa protein. Vasa is useful in the  
 XX diagnosis or treatment of conditions characterised by its aberrant  
 CC expression and/or the presence of its mutant forms. The conditions  
 CC include cancers such as biliary tract, brain, breast, colon, ovarian,  
 CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,  
 CC testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,  
 CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
 CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
 CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
 CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
 CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
 CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
 CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
 CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
 CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
 XX  
 SQ Sequence 713 AA;  
 Alignment Scores:  
 Pred. No.: 7,88e-71 Length: 713  
 Score: 82.00 Matches: 82  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.33% Indels: 0  
 DB: 22 Gaps: 0

US-09-714-865-15 (1-2172) x AAE02419 (1-713)  
 Qy 718 CAAGGACCAAAAGTGACCTACATACCCCTCTCCACCTGAGATGAGAGCTCATCTTT 777  
 Db 225 GInGlyProLyValThrTyrIleProProProProGluAspGluAspSerIlePhe 244  
 Qy 778 GCACATTATCAGACAGGACCTAAACTCGACAATAAGACACTATTCTGTGAAAGTGCT 837  
 Db 245 AlaHisTyrGlnThrGlyIleAsnPheAspLysTyrAspThrIleLeuValGluValSer 264  
 Qy 838 GGACATGATGACACACACCAATCTGACTTTTGAAGAAGCTAATCTGTGACACACTG 897  
 Db 265 GLYHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuGlySerThrLeu 284  
 Qy 898 AATTAACAACCTGCTTAACCTGTTTACTAGCTTACCTCGTGCAGAAATACACTATT 957  
 Db 285 AsnAsnAsnIleAlaLysAlaGlyTyrThrLysLeuThrProValGlnLysTyrSerIle 304  
 Qy 958 CCTATC 963  
 Db 305 Profile 306  
 RESULT 7  
 AAE02418  
 ID AAE02418 standard; Protein: 722 AA.  
 AC AAE02418;  
 XX  
 DT 10-AUG-2001 (first entry)  
 XX  
 DE Mouse vasa protein.  
 XX  
 KM Mouse; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;  
 KM pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;  
 KM thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;  
 KM melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;  
 KM choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;  
 KM myelogenous; multiple myeloma; Paget's disease; osteosarcoma;  
 KM Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
 KM leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
 KM fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
 KM teratoma; mediastinal; intracranial.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200136445-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 16-NOV-2000; 2000WO-US31485.  
 XX  
 PR 18-NOV-1999; 99US-0166394.  
 XX  
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.  
 XX  
 PI Castriillon DH;  
 XX  
 XX WPI; 2001-355606/37.  
 XX  
 DR Novel vasa polynucleotides useful in the diagnosis or treatment of  
 PT conditions characterized by aberrant expression and/or presence of  
 PT mutant forms of vasa polynucleotides or polypeptides -  
 PT  
 Claim 4; Page 54-56; 66pp; English.  
 XX  
 PS The present sequence is mouse vasa protein. Vasa is useful in the  
 XX diagnosis or treatment of conditions characterised by its aberrant  
 CC expression and/or the presence of its mutant forms. The conditions  
 CC include cancers such as biliary tract, brain, breast, colon, ovarian,  
 CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,  
 CC testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,  
 CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
 CC



CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
 CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
 CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
 CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
 CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
 CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
 CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
 CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
 CC  
 XX Sequence 722 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2,81e-53 Length: 722  
 Score: 64.00 Matches: 64  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.84% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-714-865-15 (1-2172) x AAE02418 (1-722)  
 QY 718 CAAGGACCAAAAGTACCTACATACCCCTCCCTCCAGATGAGATGAGATCCCATCTTT 777  
 Db 239 GlnGlyProLysValThrTrpIleProProProGluAspGluAspSerIlePhe 258  
 QY 778 GCACATTTATCAGACAGGACATTAACCTTCGACAAATACGACATATCTTGTGGAAGTGTCT 837  
 Db 259 AlaHisTrpGlnThrGlyIleAsnPhaAspLysTrpAspThrIleLeuValGluValSer 278  
 QY 838 GGACATGATGACACCAACGACATTTGACTTTGAAAGAGCTATCTGTGCAGACTG 897  
 Db 279 GlyHisAspAlaProProAlaIleLeuThrPheGluGluAlaAsnLeuGlyGlnThrIleu 298  
 QY 898 AATAACAGACATT 909  
 Db 299 AsnAsnAsnIle 302  
 RESULT 8  
 AAE02423  
 ID AAE02423 standard; peptide: 30 AA.  
 AC AAE02423;  
 XX  
 DT 10-AUG-2001 (first entry)  
 DE  
 XX Epitope #1 of human vasa protein, used to generate vasa antibodies.  
 XX  
 KW Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;  
 KW ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;  
 KW colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;  
 KW skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;  
 KW medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;  
 KW acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;  
 KW osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
 KW leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
 KW fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
 KW teratoma; mediastinal; intracranial.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200136445-A1.  
 PN  
 PD 25-MAY-2001.  
 PF 16-NOV-2000; 2000WO-US31485.  
 PR 18-NOV-1999; 99US-0166394.  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 XX  
 XX Caerillon DH;  
 PI  
 DR WPI; 2001-355606/37.

XX  
 PT Novel vasa polynucleotides useful in the diagnosis or treatment of  
 PT conditions characterized by aberrant expression and/or presence of  
 PT mutant forms of vasa polynucleotides or polypeptides -  
 PS Claim 19, Page 36; 66pp; English.  
 XX  
 XX The present sequence is epitope of human vasa protein that has germ cell  
 CC specific expression and is believed to play a determinative role in gonad  
 CC development. Germ cells are specialised to produce haploid gametes in  
 CC multicellular organisms. Vasa is useful in the diagnosis or treatment  
 CC of conditions characterised by its aberrant expression and/or the  
 CC presence of its mutant forms. The conditions include cancers such as  
 CC biliary tract, brain, breast, colon, ovarian, pancreas, prostate,  
 CC colorectal, oral, liver, lung, skin, basocellular, testis, renal,  
 CC thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,  
 CC melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
 CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
 CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
 CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
 CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
 CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
 CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
 CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
 CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
 XX  
 SQ Sequence 30 AA;  
 Alignment Scores:  
 Pred. No.: 5.75e-20 Length: 30  
 Score: 30.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.14% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-714-865-15 (1-2172) x AAE02423 (1-30)  
 QY 349 TGGCAGATTAATCCAGACGGAACGAGGCTTTCCAGAGAGGGGCTATCGAGATGGA 408  
 Db 1 CysGluAspAsnProThrArgAsnArgGlyPheSerLysArgGlyTrpArgAspGly 20  
 QY 409 AATAATTCAGAGCTTCAGGCGCATACAGA 438  
 Db 21 AsnAsnSerGluAlaSerGlyProTyrArg 30  
 RESULT 9  
 AAE02424  
 ID AAE02424 standard; peptide: 29 AA.  
 AC AAE02424;  
 XX  
 DT 10-AUG-2001 (first entry)  
 DE  
 XX Epitope #2 of human vasa protein, used to generate vasa antibodies.  
 XX  
 KW Human; vasa; germ cell; gonad development; therapy; cancer; oral; brain;  
 KW ovarian; biliary tract; breast; pancreas; prostate; colorectal; cervical;  
 KW colon; lung; testis; renal; thyroid; oesophageal; endometrial; gastric;  
 KW skin; liver; lymphoma; melanoma; glioblastoma; neuroblastoma; neoplasm;  
 KW medulloblastoma; choriocarcinoma; squamous cell carcinoma; leukaemia;  
 KW acute lymphocytic; myelogenous; multiple myeloma; Paget's disease;  
 KW osteosarcoma; Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
 KW leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
 KW fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
 KW teratoma; mediastinal; intracranial.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200136445-A1.  
 PN  
 PD 25-MAY-2001.  
 PA  
 XX  
 XX

XX 16-NOV-2000; 2000MO-US31485.  
XX  
XX 18-NOV-1999; 99US-0166394.  
XX  
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
XX  
XX Castillon DH;  
XX  
XX WPI, 2001-355606/37.  
XX  
XX Novel vasa polynucleotides useful in the diagnosis or treatment of  
PT conditions characterized by aberrant expression and/or presence of  
PT mutant forms of vasa polynucleotides or polypeptides -  
XX  
XX  
XX Claim 19; Page 36; 66pp; English.  
XX  
XX The present sequence is epitope of human vasa protein that has germ cell  
CC specific expression and is believed to play a determinative role in gonad  
CC Epitope is used in the generation of anti-human vasa antibodies.  
CC development. Germ cells are specialised to produce haploid gametes in  
CC multicellular organisms. Vasa is useful in the diagnosis or treatment  
CC of conditions characterised by its aberrant expression and/or the  
CC presence of its mutant forms. The conditions include cancers such as  
CC biliary tract, brain, breast, colon, ovarian, pancreas, prostate,  
CC colorectal, oral, liver, lung, skin, basocellular, testis, renal,  
CC thyroid, cervical, endometrial, oesophageal and gastric, lymphomas,  
CC melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
XX  
XX SQ Sequence 29 AA;  
XX  
XX Alignment Scores:  
XX Pred. No.: 5,45e-19 Length: 29  
XX Score: 29.00 Matches: 29  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 4.01% Indels: 0  
XX DB: 22 Gaps: 0  
XX  
XX US-09-714-865-15 (1-2172) x AAE02424 (1-29)  
XX  
XX QY 2083 GTTGATACGAGAAAGGCAAGAGCACTTGAACACAGCTGGTCTTCTTCACAGCT 2142  
XX |||||||  
XX 1 ValaapthirargulvgcllybserthrleuamnthralaaglyphaserSerSerrargala 20  
XX  
XX QY 2143 CCCAATCCAGTAGATGATGATGATG 2169  
XX |||||||  
XX 21 ProaenProvalaspaspclusertrp 29  
XX  
XX RESULT 10  
XX AAE02420  
XX ID AAE02420 standard; Protein; 700 AA.  
XX AC AAE02420;  
XX  
XX 10-AUG-2001 (first entry)  
XX  
XX Frog vasa protein.  
XX  
XX Frog; vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;  
XX pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;  
XX thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;  
XX melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;  
XX choriocarcinoma; squamous cell carcinoma; leukaemia; acute lymphocytic;  
XX myelogenous; multiple myeloma; Paget's disease; osteosarcoma;

KW Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
KW leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
KW fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
KW teratoma; mediastinal; intracranial.  
XX  
XX Xenopus laevis.  
XX  
XX WO200136445-A1.  
XX  
XX 25-MAY-2001.  
XX  
XX 16-NOV-2000; 2000MO-US31485.  
XX  
XX 18-NOV-1999; 99US-0166394.  
XX  
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
XX  
XX Castillon DH;  
XX  
XX WPI, 2001-355606/37.  
XX  
XX Novel vasa polynucleotides useful in the diagnosis or treatment of  
PT conditions characterized by aberrant expression and/or presence of  
PT mutant forms of vasa polynucleotides or polypeptides -  
XX  
XX  
XX Claim 4; Page 58-59; 66pp; English.  
XX  
XX The present sequence is frog vasa protein. Vasa is useful in the  
CC diagnosis or treatment of conditions characterised by its aberrant  
CC expression and/or the presence of its mutant forms. The conditions  
CC include cancers such as biliary tract, brain, breast, colon, ovarian,  
CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,  
CC testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,  
CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
XX  
XX SQ Sequence 700 AA;  
XX  
XX Alignment Scores:  
XX Pred. No.: 2.86e-14 Length: 700  
XX Score: 24.00 Matches: 24  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 3.31% Indels: 0  
XX DB: 22 Gaps: 0  
XX  
XX US-09-714-865-15 (1-2172) x AAE02420 (1-700)  
XX  
XX QY 970 GCAGACGAGATTGATGCTTGCTCAACAGAGCTCGGAGAGCTGCTTTCTC 1029  
XX |||||||  
XX 310 AlacllyargaspbeuMetrAlaCyAlaGlnthrclysergllystrAlaAlaIlePheIeu 329  
XX  
XX QY 1030 CTCACCAATTTTG 1041  
XX |||||||  
XX 330 LeuProIleIeu 333  
XX  
XX RESULT 11  
XX AAE02421  
XX ID AAE02421 standard; Protein; 700 AA.  
XX AC AAE02421;  
XX  
XX 10-AUG-2001 (first entry)  
XX  
XX Danilo reio vasa protein.  
XX

KM Vasa; therapy; cancer; oral; brain; ovarian; biliary tract; lung;  
 KM pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;  
 KM thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;  
 KM melanoma; glioblastoma; neuroblastoma; neoplasia; medulloblastoma;  
 KM choriocarcinoma; squamous cell carcinoma; leukemia; acute lymphocytic;  
 KM myelogenous; multiple myeloma; Paget's disease; osteosarcoma;  
 KM Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
 KM leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
 KM fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
 KM teratoma; mediastinal; intracranial.  
 XX  
 OS Danio reio.  
 XX  
 PN WO200136445-A1.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 16-NOV-2000; 2000WO-US31485.  
 XX  
 PR 18-NOV-1999; 99US-0166394.  
 XX  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 XX  
 PI Caet'illon Df;  
 XX  
 DR WPI; 2001-355606/37.  
 XX  
 PT Novel vasa polynucleotides useful in the diagnosis or treatment of  
 PT conditions characterized by aberrant expression and/or presence of  
 PT mutant forms of vasa polynucleotides or polypeptides -  
 XX  
 PS Claim 4; Page 59-61; 66pp; English.  
 XX  
 CC The present sequence is Danio reio vasa protein. Vasa is useful in the  
 CC diagnosis or treatment of conditions characterised by its aberrant  
 CC expression and/or the presence of its mutant forms. The conditions  
 CC include cancers such as biliary tract, brain, breast, colon, ovarian,  
 CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,  
 CC testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,  
 CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
 CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
 CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
 CC acquired immune deficiency syndrome (AIDS) associated leukaemias,  
 CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
 CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
 CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
 CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
 CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
 CC  
 XX  
 SQ Sequence 700 AA;  
 Alignment Scores:  
 Pred. No.: 2,86e-14 Length: 700  
 Score: 24.00 Matches: 24  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.31% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-714-865-15 (1-2172) x AAE02421 (1-700)  
 QY 970 GCAGGAGGATTTGATGCTTGTCTCAACAGGCTTGGGAAGACTGGCGCTTTTCTC 1029  
 Db 297 AlaaIyArGaPpLeuMeAlaCyAlaGlnThrcIySerGlyVsrThraAlaPheLeu 316  
 QY 1030 CTACCAATTTTG 1041  
 Db 317 LeuProlleu 320  
 RESULT 12  
 ABB59954 standard; Protein; 661 AA.  
 ID ABB59954  
 XX

AC ABB59954;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 6654.  
 XX  
 KM Drosophila; developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 XX  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li FWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL04057.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 6654; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins  
 CC (ABBS7737-ABBS72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcr\_sequences.  
 CC  
 XX  
 SQ Sequence 661 AA;  
 Alignment Scores:  
 Pred. No.: 2,72e-13 Length: 661  
 Score: 23.00 Matches: 23  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.18% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-714-865-15 (1-2172) x ABB59954 (1-661)  
 QY 973 GGAGGAGATTTGATGCTTGTCTCAACAGGCTTGGGAAGACTGGCGCTTTTCTCTA 1032  
 Db 282 GIyArGaPpLeuMeAlaCyAlaGlnThrcIySerGlyVsrThraAlaPheLeu 301  
 QY 1033 CCAATTTTG 1041  
 Db 302 Prolleu 304  
 RESULT 13  
 AAE02422  
 ID AAE02422 standard; Protein; 661 AA.  
 AC AAE02422;  
 XX  
 AC AAE02422;  
 XX  
 DT 10-AUG-2001 (first entry)  
 XX

DE Fruit fly vasa protein.  
XX  
KM Vasa; therapy: cancer; oral; brain; ovarian; biliary tract; lung;  
KM pancreas; prostate; colorectal; cervical; breast; colon; testis; renal;  
KM thyroid; oesophageal; endometrial; gastric; skin; liver; lymphoma;  
KM melanoma; glioblastoma; neuroblastoma; neoplasm; medulloblastoma;  
KM choriocarcinoma; squamous cell carcinoma; leukemia; acute lymphocytic;  
KM myelogenous; multiple myeloma; Paget's disease; osteosarcoma;  
KM Acquired immune deficiency syndrome; AIDS; Bowen's disease;  
KM leiomyosarcoma; rhabdomyosarcoma; liposarcoma; Kaposi's sarcoma;  
KM fibrosarcoma; tumour; testicular; seminoma; ovarian; dysgerminoma;  
KM teratoma; mediastinal; intracranial; fruit fly.  
OS Drosophila melanogaster.  
XX  
PN WO200136445-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 16-NOV-2000; 2000WO-US31485.  
XX  
PR 18-NOV-1999; 99US-0166394.  
XX  
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
PI Castrixllon DH;  
XX  
DR WPI; 2001-355606/37.  
XX  
PT Novel vasa polynucleotides useful in the diagnosis or treatment of  
PT conditions characterized by aberrant expression and/or presence of  
PT mutant forms of vasa polynucleotides or polypeptides -  
XX  
PS Claim 4; Page 61-62; 66pp; English.  
XX  
CC The present sequence is fruit fly vasa protein. Vasa is useful in the  
CC diagnosis or treatment of conditions characterised by its aberrant  
CC expression and/or the presence of its mutant forms. The conditions  
CC include cancers such as biliary tract, brain, breast, colon, ovarian,  
CC pancreas, prostate, colorectal, oral, liver, lung, skin, basocellular,  
CC testis, renal, thyroid, cervical, endometrial, oesophageal and gastric,  
CC lymphomas, melanomas, glioblastomas, neuroblastomas, medulloblastomas,  
CC choriocarcinoma, squamous cell carcinoma, haematological neoplasms,  
CC acute lymphocytic and myelogenous leukaemia, multiple myeloma,  
CC Acquired immune deficiency syndrome (AIDS) associated leukaemias,  
CC intraepithelial neoplasms, Bowen's disease, Paget's disease, sarcomas  
CC such as leiomyosarcoma, rhabdomyosarcoma, liposarcoma, fibrosarcoma,  
CC Kaposi's sarcoma and osteosarcoma, tumours such as testicular tumour  
CC (eg. seminoma), ovarian tumour (eg. dysgerminoma or teratoma) and tumour  
CC of an extragonadal tissue (eg. mediastinal or an intracranial tumour).  
XX  
SQ Sequence 661 AA;  
Alignment Scores:  
Pred. No.: 2.72e-13 Length: 661  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.18% Indels: 0  
Gaps: 0  
US-09-714-865-15 (1-2172) x AAE02422 (1-661)  
OY 973 GGACGAGATTGATGGCTTGGCTCAACAGAGGTCTGGGAGACCTGGCGCTTTCTCTTA 1032  
DB 282 GtYAgAaPLeMwEclACyAlAGInThrgISerGlySThrlAlAlaPhenLeu 301  
OY 1033 CCAATTG 1041  
DB 302 Prolleu 304  
RESULT 14  
AAB38176

ID AAB38176 standard; Protein; 112 AA.  
XX  
AC AAB38176;  
XX  
DT 30-JAN-2001 (first entry)  
XX  
DE Gene 6 human secreted protein homologous amino acid sequence #115.  
XX  
KM Human; secreted protein; diagnosis; immunosuppressive; antirhectic;  
KM antineumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
KM cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;  
KM fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
KM hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
KM cerebrovascular disorder; nervous system disorder; ocular disorder;  
KM wound healing; skin aging; food additive; preservative.  
XX  
OS Homo sapiens.  
XX  
PN WO200058468-A2.  
XX  
PD 05-OCT-2000.  
XX  
PF 22-MAR-2000; 2000WO-US07526.  
XX  
PR 26-MAR-1999; 99US-0126600.  
XX  
PR 22-DEC-1999; 99US-0171550.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
DR WPI; 2000-611713/58.  
XX  
PT Nucleic acids encoding human secreted proteins. used to prevent, treat,  
PT ameliorate, or diagnose conditions such as autoimmune disorders, skin  
PT disorders and cancer -  
XX  
PS Disclosure; Page 363; 374pp; English.  
XX  
CC The polynucleotide sequences given in AAC69399 to AAC69445 encode the  
CC human secreted proteins given in AAB38119 to AAB38165. AAB38166 to  
CC AAB38201 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissues  
CC and cells the genes are expressed in. Example of activities include:  
CC immunosuppressive; antirhectic; antineumatic; antiproliferative;  
CC cyostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
CC neuroprotective; antibacterial; virucide; fungicide; and  
CC ophthalmological. The polynucleotides and polypeptides can be used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases, hyperproliferative disorders, cardiovascular  
CC disorders, cerebrovascular disorders, angiogenesis, nervous system  
CC disorders, infections caused by bacteria, viruses and fungi and ocular  
CC disorders. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. AAC69390 to AAC69398 and  
CC AAB38118 represent sequences used in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 112 AA;  
XX  
Alignment Scores:  
Pred. No.: 3.14e-12 Length: 112  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.04% Indels: 0

DB: 21 Gaps: 0  
US-09-714-865-15 (1-2172) x AAB38176 (1-112)  
QY 976 CGAGATTGATGGCTTGCTCAAAACAGGGCTTGGGAAGACTGGGGCTTTCTCTACCA 1035  
71 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyThrAlaAlaPheLeuLeuPro 90  
Db 71 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyThrAlaAlaPheLeuLeuPro 90  
QY 1036 ATTTTG 1041  
91 IleLeu 92  
Db 91 IleLeu 92  
RESULT 15  
AAB38177  
ID AAB38177 standard; Protein; 112 AA.  
XX  
AC AAB38177;  
XX  
XX 30-JAN-2001 (first entry)  
XX  
XX Human secreted protein sequence encoded by gene 6 SEQ ID NO:116.  
XX  
XX Human: secreted protein; diagnosis; immunosuppressive; antiarthritic;  
XX antiinematic; antiproliferative; cyostatic; cardiant; vasotrophic;  
XX cerebroprotective; nootropic; neuroprotective; antibacterial; viruicide;  
XX fungicide; ophthalmological; gene therapy; autoimmune disease; infection;  
XX hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
XX cerebrovascular disorder; nervous system disorder; ocular disorder;  
XX wound healing; skin aging; food additive; preservative.  
XX  
XX Homo sapiens.  
XX  
XX WO200058468-A2.  
XX  
XX 05-OCT-2000.  
XX  
XX 22-MAR-2000; 2000WO-US07526.  
XX  
XX 26-MAR-1999; 99US-012600.  
XX  
XX 22-DEC-1999; 99US-0171550.  
XX  
XX (HUMA-) HUMAN GENOME SCT INC.  
XX  
XX Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI; 2000-611713/58.  
XX  
XX Nucleic acids encoding human secreted proteins, used to prevent, treat,  
XX ameliorate, or diagnose conditions such as autoimmune disorders, skin  
XX disorders and cancer -  
XX  
XX  
XX Disclosure; Page 363-364; 374pp; English.  
XX  
XX The polynucleotide sequences given in AAC69399 to AAC69445 encode the  
XX human secreted proteins given in AAB38119 to AAB38165. AAB38166 to  
XX AAB38201 represent human secreted polypeptide sequences and proteins  
XX homologous to them, which are given in the exemplification of the present  
XX invention. Human secreted proteins have activities based on the tissues  
XX and cells the genes are expressed in. Example of activities include:  
XX immunosuppressive; antiarthritic; antiinematic; antiproliferative;  
XX cyostatic; cardiant; vasotrophic; cerebroprotective; nootropic;  
XX neuroprotective; antibacterial; viruicide; fungicide; and  
XX ophthalmological. The polynucleotides and polypeptides can be are used to  
XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
XX rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
XX in diagnosing a pathological condition or susceptibility to a  
XX pathological condition. Disorders which are diagnosed or treated include  
XX autoimmune diseases, hyperproliferative disorders, cardiovascular  
XX disorders, cerebrovascular disorders, angiogenesis, nervous system  
XX disorders, infections caused by bacteria, viruses and fungi and ocular  
XX disorders. The polypeptides can also be used to aid wound healing and  
XX epithelial cell proliferation, to prevent skin aging due to sunburn, to  
XX maintain organs before transplantation, for supporting cell culture of

CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. AAC69390 to AAC69398 and  
CC AAB38118 represent sequences used in the exemplification of the present  
CC invention.

XX Sequence 112 AA;

QY Alignment Scores:

Pred. No.: 3,14e-12 Length: 112

Score: 22.00 Matches: 22

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.04% Indels: 0

DB: 21 Gaps: 0

US-09-714-865-15 (1-2172) x AAB38177 (1-112)

QY 976 CGAGATTGATGGCTTGCTCAAAACAGGGCTTGGGAAGACTGGGGCTTTCTCTACCA 1035

Db 71 ArgAspLeuMetAlaCysAlaGlnThrGlySerGlyThrAlaAlaPheLeuLeuPro 90

QY 1036 ATTTTG 1041

Db 91 IleLeu 92

Search completed: June 10, 2003, 17:04:54  
Job time : 82.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 11:55:35 ; Search time 2772 Seconds  
(without alignments)  
12689.970 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 2172  
Sequence: 1 atcgsggagatgaagatcgga.....tagatgatgcacatcgat 2172

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 1000 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estnu: \*  
5: em\_estoy: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gse: \*  
18: em\_gse\_hum: \*  
19: em\_gse\_inv: \*  
20: em\_gse\_pin: \*  
21: em\_gse\_vrt: \*  
22: em\_gse\_fun: \*  
23: em\_gse\_mam: \*  
24: em\_gse\_mus: \*  
25: em\_gse\_oher: \*  
26: em\_gse\_pro: \*  
27: em\_gse\_rod: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806	83.1	2523	11	BC030638 Homo sapi
2	616	28.4	682	12	BG1718234
3	616	28.4	683	13	B1560849
4	612	28.2	798	12	BG1717110
5	535	24.6	792	14	BQ429126
6	509	23.4	747	12	BG1719449

7	480	22.1	657	12	BG171812
8	469	21.6	651	12	BG1717354
9	441	20.3	822	13	B1464267
10	440	20.3	846	12	BG1718634
11	430	19.8	861	13	B1461393
12	422	19.4	830	13	B1462207
13	413	19.0	769	12	BG171822
14	412	19.0	743	13	B1463502
15	392	18.0	841	13	B1459115
16	381	17.5	884	13	B1560499
17	361	16.6	918	13	B1561001
18	357	15.5	810	13	B1463846
19	328	15.1	758	13	B1559479
20	321	14.8	852	9	AL042306
21	277	12.8	585	9	AA399611
22	267	12.3	847	13	B1562624
23	246	11.3	555	9	AI217144
24	246	11.3	917	13	B1463447
25	233	10.7	335	9	AA383535
26	219	10.1	491	9	AA398976
27	204	9.4	700	12	BG1718861
28	202	9.3	347	9	AI953070
29	161	7.4	604	13	B1460628
30	132	6.1	300	9	AI337133
31	132	6.1	300	9	AI654417
32	132	6.1	300	10	BE550307
33	132	6.1	300	10	BE550448
34	114	5.2	360	10	AA449170
35	108	5.0	368	10	AA469018
36	64	2.9	371	9	AA758412
37	63	2.9	469	9	AI025074
38	58	2.7	2759	11	AK014844
39	53	2.4	374	17	BE3809
40	49	2.3	103	9	AA400066
41	44	2.0	546	13	BMS39751
42	42	1.9	665	17	A2648249
43	41	1.9	527	17	A2806430
44	41	1.9	589	17	AO932766
45	39	1.8	79	9	AA401568
46	39	1.8	100	9	AA862553
47	36	1.7	772	13	B1830532
48	34	1.6	637	13	B1461238
49	33	1.5	670	14	BQ443738
50	28	1.3	420	17	AO543635
51	26	1.2	430	9	AI843771
52	25	1.2	737	17	AG123731
53	23	1.1	141	12	BE384091
54	23	1.1	346	10	BE192466
55	23	1.1	355	9	AA087515
56	23	1.1	435	10	BE227079
57	23	1.1	441	9	AA517647
58	23	1.1	466	13	B1190619
59	23	1.1	497	12	BF150205
60	23	1.1	518	13	BM125854
61	23	1.1	555	17	A2401250
62	23	1.1	587	9	AA896298
63	23	1.1	600	12	BG804318
64	23	1.1	600	12	BG808188
65	23	1.1	610	10	BE382075
66	23	1.1	621	9	AA098277
67	23	1.1	666	12	BF452294
68	23	1.1	696	14	BQ265507
69	23	1.1	714	12	BF166660
70	23	1.1	734	12	BF166601
71	23	1.1	753	13	B1689003
72	23	1.1	762	13	BG169003
73	23	1.1	790	13	B1078164
74	23	1.1	824	12	BE916016
75	23	1.1	847	12	BG519330
76	23	1.1	855	12	BE607262
77	23	1.1	864	12	BG293320
78	23	1.1	891	9	AI324427
79	23	1.1	906	14	BQ964084

80	23	1.1	941	14	BQ886382	BQ886382 AGENCOURT	C 153	20	0.9	548	13	BM376181	BM376181 EBma01_SQ
81	23	1.1	959	14	BQ714521	BQ714521 AGENCOURT	C 154	20	0.9	549	13	BM324340	BM324340 BJ324340
82	23	1.1	965	14	BQ714287	BQ714287 AGENCOURT	C 155	20	0.9	552	10	AM030845	AM030845 EST374135
83	23	1.1	969	14	BQ288214	BQ288214 AGENCOURT	C 156	20	0.9	552	10	BE342512	BE342512 EST395356
84	23	1.1	1099	12	BE913621	BE913621 60167566	C 157	20	0.9	559	14	BQ460503	BQ460503 HA09G037
85	22	1.0	604	17	A2448672	A2448672 1M0246D03	C 158	20	0.9	561	17	A2348237	A2348237 1M0084K18
86	22	1.0	635	12	BG448350	BG448350 NF070H07E	C 159	20	0.9	566	10	AM830213	AM830213 sm24b07.y
87	22	1.0	727	12	BG645837	BG645837 EST507446	C 160	20	0.9	566	13	BM328081	BM328081 BJ328081
88	21	1.0	255	9	AV125950	AV125950 ddd22b08	C 161	20	0.9	566	13	BM368848	BM368848 EBma08_SQ
89	21	1.0	255	12	BG512339	BG512339 C1TB1-El-	C 162	20	0.9	567	10	AM567872	AM567872 g166G05.y
90	21	1.0	530	17	AQ889388	AQ889388 C1TB1-El-	C 163	20	0.9	582	13	BM369548	BM369548 EBem07_SQ
91	21	1.0	588	14	BQ780602	BQ780602 UI-R-F0-	C 164	20	0.9	582	13	BM372228	BM372228 EBic05_SQ
92	21	1.0	693	17	AQ854670	AQ854670 CPG2132A	C 165	20	0.9	582	13	BM376771	BM376771 EBem05_SQ
93	21	1.0	780	10	AM153705	AM153705 f124G05.y	C 166	20	0.9	585	17	A2906662	A2906662 RPCI-24-2
94	21	1.0	860	9	A1049281	A1049281 uc85f11.y	C 167	20	0.9	587	14	BQ460364	BQ460364 HA09O16-
95	21	1.0	860	9	BQ891158	BQ891158 AGENCOURT	C 168	20	0.9	594	13	BM348795	BM348795 BJ448795
96	21	1.0	1007	17	CNS03B3C	AL236316 Teiraodon	C 169	20	0.9	595	10	BE190477	BE190477 s018a02.y
97	20	0.9	190	14	BQ756631	BQ756631 EBem09_SQ	C 170	20	0.9	598	10	AV922083	AV922083 AV922083
98	20	0.9	247	9	AV274771	AV274771 AV274771	C 171	20	0.9	599	13	BM377572	BM377572 EBem04_SQ
99	20	0.9	247	10	BM152670	BM152670 BM152670	C 172	20	0.9	608	10	AV916786	AV916786 AV916786
100	20	0.9	282	9	A1795264	A1795264 RRMCAC966	C 173	20	0.9	610	10	BE555612	BE555612 sg04f04.y
101	20	0.9	300	9	AJ464817	AJ464817 AJ464817	C 174	20	0.9	610	13	BM461931	BM461931 BJ461931
102	20	0.9	300	9	AJ472851	AJ472851 AJ472851	C 175	20	0.9	613	13	BM527230	BM527230 ba151c11
103	20	0.9	327	14	BQ458661	BQ458661 HA04Q08x	C 176	20	0.9	616	13	BM528838	BM528838 BJ528838
104	20	0.9	357	13	BM371273	BM371273 EBma08_SQ	C 177	20	0.9	620	13	BM530501	BM530501 BJ530501
105	20	0.9	360	9	AJ460815	AJ460815 AJ460815	C 178	20	0.9	623	13	BM780201	BM780201 EBma08_SQ
106	20	0.9	360	9	AJ464819	AJ464819 AJ464819	C 179	20	0.9	624	17	AO566420	AO566420 HS_209F.A
107	20	0.9	360	9	AJ481941	AJ481941 AJ481941	C 180	20	0.9	629	13	BM177213	BM177213 EST518158
108	20	0.9	360	14	C39342	C39342 Yuj1	C 181	20	0.9	632	9	AL785068	AL785068 AL785068
109	20	0.9	374	13	BM096712	BM096712 EBma07_SQ	C 182	20	0.9	633	9	AJ433088	AJ433088 AL785068
110	20	0.9	382	12	BG405848	BG405848 ea3c3f03	C 183	20	0.9	644	12	BG044371	BG044371 saa27f05
111	20	0.9	394	13	BM373875	BM373875 EBma03_SQ	C 184	20	0.9	644	14	BQ470988	BQ470988 HX04K05x
112	20	0.9	397	13	BM444825	BM444825 EBem09_SQ	C 185	20	0.9	645	14	BM793997	BM793997 EST_2935
113	20	0.9	398	12	BE76776	BE76776 MY-20-C-0	C 186	20	0.9	650	14	BM768604	BM768604 EBx008_SQ
114	20	0.9	399	17	AQ074478	AQ074478 C1T-HSP-2	C 187	20	0.9	660	17	AQ264780	AQ264780 C1TB1-El-
115	20	0.9	412	13	BM098910	BM098910 EBP105_SQ	C 188	20	0.9	663	14	BQ768464	BQ768464 EBic08_SQ
116	20	0.9	420	9	AJ460816	AJ460816 AJ460816	C 189	20	0.9	671	12	BG418184	BG418184 HVSMEK002
117	20	0.9	420	9	AJ464815	AJ464815 AJ464815	C 190	20	0.9	672	12	BF263745	BF263745 HV_CEA000
118	20	0.9	420	9	AJ464816	AJ464816 AJ464816	C 191	20	0.9	673	9	A1491087	A1491087 EST241796
119	20	0.9	420	9	AJ464818	AJ464818 AJ464818	C 192	20	0.9	676	14	BM760071	BM760071 EBP107_SQ
120	20	0.9	420	9	AJ471104	AJ471104 AJ471104	C 193	20	0.9	677	17	AQ287739	AQ287739 HX04K051A
121	20	0.9	420	9	AJ481943	AJ481943 AJ481943	C 194	20	0.9	681	14	BQ510476	BQ510476 EST617891
122	20	0.9	435	13	BM097625	BM097625 EBem04_SQ	C 195	20	0.9	684	13	BM1934953	BM1934953 EST54842
123	20	0.9	438	13	BM779839	BM779839 EBem07_SQ	C 196	20	0.9	686	13	BM393334	BM393334 500711-2-9
124	20	0.9	438	13	BM772690	BM772690 ea181f01	C 197	20	0.9	686	13	BM395422	BM395422 50072-2-9
125	20	0.9	442	10	BM785361	BM785361 BB785361	C 198	20	0.9	687	9	AU060736	AU060736 AU060736
126	20	0.9	444	9	AA874508	AA874508 vx03b10.t	C 199	20	0.9	687	14	BM764827	BM764827 EBca01_SQ
127	20	0.9	444	12	BF263744	BF263744 HV_CEA000	C 200	20	0.9	694	14	BQ542660	BQ542660 BJ542660
128	20	0.9	446	13	BM498518	BM498518 ea115c01	C 201	20	0.9	694	14	BQ509337	BQ509337 EST616752
129	20	0.9	448	12	BE750660	BE750660 202135_MA	C 202	20	0.9	700	12	BF621942	BF621942 HVSMEa000
130	20	0.9	457	13	BM7073747	BM7073747 BU073747	C 203	20	0.9	700	17	BM470168	BM470168 BHF621942
131	20	0.9	460	12	BE750662	BE750662 202137_MA	C 204	20	0.9	701	12	BF622691	BF622691 BHF621942
132	20	0.9	460	17	CNS06KXB	AL421137 T3 end of	C 205	20	0.9	702	13	BM175851	BM175851 EST517034
133	20	0.9	461	14	BQ756838	BQ756838 EBem09_SQ	C 206	20	0.9	722	17	AG105109	AG105109 Pan_17c031
134	20	0.9	462	13	BM441169	BM441169 EBed02_SQ	C 207	20	0.9	730	12	BF611759	BF611759 dB68d02.y
135	20	0.9	474	13	BM777934	BM777934 EBx007_SQ	C 208	20	0.9	732	13	BM541123	BM541123 BJ541123
136	20	0.9	477	17	AQ415685	AQ415685 RPCI-11-2	C 209	20	0.9	824	17	CNS04FMG	AL2888529 Teiraodon
137	20	0.9	480	9	AJ471105	AJ471105 AJ471105	C 210	20	0.9	825	17	A2535469	A2535469 ENTOR48TF
138	20	0.9	480	9	AJ471106	AJ471106 AJ471106	C 211	20	0.9	852	17	A2669540	A2669540 ENTLM40TR
139	20	0.9	480	9	AJ471108	AJ471108 AJ471108	C 212	20	0.9	898	17	A2541284	A2541284 ENTGF32TF
140	20	0.9	480	9	AJ481942	AJ481942 AJ481942	C 213	20	0.9	962	17	A2691214	A2691214 ENTTC88TF
141	20	0.9	480	9	AJ481944	AJ481944 AJ481944	C 214	20	0.9	983	10	BM610861	BM610861 BB610861
142	20	0.9	480	9	AJ481945	AJ481945 AJ481945	C 215	20	0.9	1026	17	CNS0783A	AL433484 T7 end of
143	20	0.9	505	13	BM947504	BM947504 HVSME1000	C 216	20	0.9	111	13	BM387357	BM387357 BJ387357
144	20	0.9	514	10	AV670989	AV670989 AV670989	C 217	20	0.9	153	9	A1926243	A1926243 w041G07.x
145	20	0.9	521	14	BQ458499	BQ458499 EBx008_SQ	C 218	20	0.9	156	9	A1046086	A1046086 TENU1172
146	20	0.9	528	9	A1929849	A1929849 HA05A22x	C 219	20	0.9	236	9	AA926895	AA926895 cm25f03..s
147	20	0.9	530	10	BE603260	BE603260 HVSMEH010	C 220	20	0.9	237	10	BE077508	BE077508 RCI-1B7060
148	20	0.9	530	13	BM387354	BM387354 BM387354	C 221	20	0.9	244	12	BM626723	BM626723 CC-66f1c1
149	20	0.9	537	13	BQ520442	BQ520442 NISC_n104	C 222	20	0.9	244	14	N31074	N31074 yx3e03.t1
150	20	0.9	540	9	AJ471107	AJ471107 AJ471107	C 223	20	0.9	248	10	BM046684	BM046684 BB046684
151	20	0.9	546	10	AW755646	AW755646 g106d02.y	C 224	20	0.9	256	10	AV361686	AV361686 AV361686
152	20	0.9	546	10	AW755646	AW755646 g106d02.y	C 225	20	0.9	258	17	A2459411	A2459411 1M0264D14





C 372	19	0.9	660	10	AM685848	AM685848	NFO30H09N	C 445	19	0.9	928	12	BE898612	BE898612	601681550
C 373	19	0.9	661	17	AZ600454	AZ600454	IM0418117	C 446	19	0.9	936	14	BO680001	BO680001	AGENCOURT
C 374	19	0.9	663	10	BB618815	BB618815	BB618815	C 447	19	0.9	936	17	CNS0310H	CNS0310H	AL247130
C 375	19	0.9	663	12	BF645868	BF645868	NFO32F08E	C 448	19	0.9	949	11	AK020377	AK020377	Mus muscu
C 376	19	0.9	663	12	BF166736	BF166736	601776251	C 449	19	0.9	961	14	BO678258	BO678258	AGENCOURT
C 377	19	0.9	665	14	BO407620	BO407620	Gd_Ed000	C 450	19	0.9	962	14	BO723669	BO723669	AGENCOURT
C 378	19	0.9	668	17	AL505041	AL505041	AL505041	C 451	19	0.9	966	12	BG484130	BG484130	602504744
C 379	19	0.9	668	17	AQ475616	AQ475616	CTBTI-ET-1	C 452	19	0.9	1003	13	BM471017	BM471017	AGENCOURT
C 380	19	0.9	674	17	AZ417760	AZ417760	IM0193M06	C 453	19	0.9	1007	17	CNS04ADN	CNS04ADN	Tetraodon
C 381	19	0.9	677	14	BM283395	BM283395	K-EST0101	C 454	19	0.9	1013	12	BE899378	BE899378	601681309
C 382	19	0.9	680	12	BF645004	BF645004	NFO20F10E	C 455	19	0.9	1018	13	BM462593	BM462593	AGENCOURT
C 383	19	0.9	680	12	BE890333	BE890333	601431788	C 456	19	0.9	1027	17	CNS0310E	CNS0310E	Tetraodon
C 384	19	0.9	680	13	BI262816	BI262816	NFO91G06E	C 457	19	0.9	1040	14	BM904892	BM904892	Tetraodon
C 385	19	0.9	682	17	AG042230	AG042230	Pan trogl	C 458	19	0.9	1055	17	CNS01NWP	CNS01NWP	Tetraodon
C 386	19	0.9	685	10	BE265000	BE265000	601193829	C 459	19	0.9	1069	17	CNS02TNP	CNS02TNP	Tetraodon
C 387	19	0.9	686	13	BJ143867	BJ143867	BJ143867	C 460	19	0.9	1082	13	BM541646	BM541646	AGENCOURT
C 388	19	0.9	687	12	BF643529	BF643529	NFO21F11E	C 461	19	0.9	1082	13	BM806554	BM806554	AGENCOURT
C 389	19	0.9	690	17	AZ854117	AZ854117	ZM0157A04	C 462	19	0.9	1089	14	BO062618	BO062618	AGENCOURT
C 390	19	0.9	691	12	BE894491	BE894491	601433091	C 463	19	0.9	1090	13	BM550361	BM550361	AGENCOURT
C 391	19	0.9	693	13	BM393529	BM393529	50072-2-1	C 464	19	0.9	1101	17	CNS05SNT	CNS05SNT	Tetraodon
C 392	19	0.9	700	9	AL503450	AL503450	AL503450	C 465	19	0.9	1126	12	BM258475	BM258475	602379954
C 393	19	0.9	700	9	AL504817	AL504817	AL504817	C 466	19	0.9	1183	13	BM546059	BM546059	AGENCOURT
C 394	19	0.9	700	12	BG592053	BG592053	EST499895	C 467	19	0.9	1257	12	BG750045	BG750045	AGENCOURT
C 395	19	0.9	703	12	BG632720	BG632720	GH12767.3	C 468	19	0.9	1452	10	BE421594	BE421594	AGENCOURT
C 396	19	0.9	708	17	AZ208374	AZ208374	SP_0150_A	C 469	19	0.9	1671	11	BC006489	BC006489	Hom sapi
C 397	19	0.9	717	9	AA986987	AA986987	uc80h06.Y	C 470	19	0.9	2714	11	AF077033	AF077033	Hom sapi
C 398	19	0.9	717	14	BO180444	BO180444	uc80h06.Y	C 471	19	0.9	2756	11	AK010281	AK010281	Hom sapi
C 399	19	0.9	718	9	AL652289	AL652289	AL652289	C 472	19	0.9	4626	11	EC007668	EC007668	Hom sapi
C 400	19	0.9	727	17	AQ051089	AQ051089	IM010050A	C 473	19	0.9	91	9	AI032269	AI032269	OT97605.8
C 401	19	0.9	736	17	BM543923	BM543923	BOGEF30TR	C 474	19	0.9	110	10	BE319457	BE319457	NFO18G10R
C 402	19	0.9	744	13	BI102865	BI102865	602888459	C 475	19	0.9	142	13	BI1883146	BI1883146	603003.0
C 403	19	0.9	744	13	BI102865	BI102865	602888459	C 476	19	0.9	143	13	BI179433	BI179433	603003.0
C 404	19	0.9	752	12	BF046892	BF046892	EST989.Ma	C 477	19	0.9	146	13	BI18187	BI18187	603003.0
C 405	19	0.9	754	13	BI550532	BI550532	603195457	C 478	19	0.9	150	13	BG956149	BG956149	603003.0
C 406	19	0.9	756	12	BG599810	BG599810	EST504705	C 479	19	0.9	163	13	BI18074	BI18074	603003.0
C 407	19	0.9	760	17	AQ0935849	AQ0935849	CPG26858	C 480	19	0.9	171	12	BE836343	BE836343	603003.0
C 408	19	0.9	760	17	BM307225	BM307225	RPCI-24-2	C 481	19	0.9	171	13	BI17964	BI17964	603003.0
C 409	19	0.9	762	12	BF302156	BF302156	602033110	C 482	19	0.9	180	9	AJ462330	AJ462330	603003.0
C 410	19	0.9	762	12	BF302156	BF302156	602033110	C 483	19	0.9	180	9	AM126455	AM126455	614072002
C 411	19	0.9	777	13	BI918073	BI918073	603181963	C 484	19	0.9	194	13	BI841305	BI841305	603003.0
C 412	19	0.9	779	17	CNS072M1	CNS072M1	clone BAO	C 485	19	0.9	204	14	BO092426	BO092426	603003.0
C 413	19	0.9	792	11	AK015776	AK015776	Mus muscu	C 486	19	0.9	207	12	BF091400	BF091400	603003.0
C 414	19	0.9	804	12	BG645554	BG645554	EST507173	C 487	19	0.9	207	13	BI472568	BI472568	603003.0
C 415	19	0.9	804	12	BG519508	BG519508	602577783	C 488	19	0.9	209	17	AZ633227	AZ633227	603003.0
C 416	19	0.9	804	14	BO115189	BO115189	EST600765	C 489	19	0.9	211	9	AV165117	AV165117	603003.0
C 417	19	0.9	805	10	BE412836	BE412836	MCG008.80	C 490	19	0.9	211	10	BE064380	BE064380	603003.0
C 418	19	0.9	814	13	BI081767	BI081767	602879661	C 491	19	0.9	211	12	BG135137	BG135137	603003.0
C 419	19	0.9	821	9	AU132682	AU132682	AU132682	C 492	19	0.9	212	13	BI429942	BI429942	603003.0
C 420	19	0.9	821	17	AQ050907	AQ050907	IM010004C	C 493	19	0.9	213	14	BO077929	BO077929	603003.0
C 421	19	0.9	822	17	CNS035M4	CNS035M4	Tetraodon	C 494	19	0.9	215	14	BO092923	BO092923	603003.0
C 422	19	0.9	822	13	BI824613	BI824613	603033546	C 495	19	0.9	221	13	BI707873	BI707873	603003.0
C 423	19	0.9	832	12	BE865790	BE865790	601678176	C 496	19	0.9	233	17	AZ204821	AZ204821	603003.0
C 424	19	0.9	835	17	AZ684287	AZ684287	ENTL164TF	C 497	19	0.9	238	10	AM213648	AM213648	603003.0
C 425	19	0.9	837	9	AU142864	AU142864	AU142864	C 498	19	0.9	239	17	AZ481633	AZ481633	603003.0
C 426	19	0.9	843	10	BE640750	BE640750	Cr12.1.F1	C 499	19	0.9	240	9	AU071907	AU071907	603003.0
C 427	19	0.9	843	17	AQ576114	AQ576114	IM0100887	C 500	19	0.9	244	10	AM418170	AM418170	603003.0
C 428	19	0.9	856	17	AZ205932	AZ205932	SP_0105.A	C 501	19	0.9	244	13	AV363308	AV363308	603003.0
C 429	19	0.9	856	17	CNS06DVO	CNS06DVO	602368063	C 502	19	0.9	248	10	BI882897	BI882897	603003.0
C 430	19	0.9	861	17	CNS06DVO	CNS06DVO	602368063	C 503	19	0.9	248	17	AQ412245	AQ412245	603003.0
C 431	19	0.9	866	14	BO723018	BO723018	71_end_of	C 504	19	0.9	249	9	BE528509	BE528509	603003.0
C 432	19	0.9	873	17	CNS03005	CNS03005	AL252734	C 505	19	0.9	250	9	AA605634	AA605634	603003.0
C 433	19	0.9	877	17	CNS06KXW	CNS06KXW	77_end_of	C 506	19	0.9	254	13	BM544125	BM544125	603003.0
C 434	19	0.9	880	12	BG255707	BG255707	602368063	C 507	19	0.9	254	13	BI18933	BI18933	603003.0
C 435	19	0.9	883	17	CNS0320F	CNS0320F	Tetraodon	C 508	19	0.9	254	13	BI429596	BI429596	603003.0
C 436	19	0.9	886	14	BO936221	BO936221	AGENCOURT	C 509	19	0.9	255	13	BI182880	BI182880	603003.0
C 437	19	0.9	890	12	BE884173	BE884173	601505928	C 510	19	0.9	255	13	BI18893	BI18893	603003.0
C 438	19	0.9	895	13	BI819955	BI819955	603035314	C 511	19	0.9	255	13	BI429543	BI429543	603003.0
C 439	19	0.9	900	14	BO899473	BO899473	AGENCOURT	C 512	19	0.9	256	17	BM533459	BM533459	603003.0
C 440	19	0.9	906	14	BO218557	BO218557	AGENCOURT	C 513	19	0.9	256	17	BM502507	BM502507	603003.0
C 441	19	0.9	918	12	BF167886	BF167886	601775619	C 514	19	0.9	258	13	BI139006	BI139006	603003.0
C 442	19	0.9	918	17	CNS03957	CNS03957	Tetraodon	C 515	19	0.9	261	13	BI474332	BI474332	603003.0
C 443	19	0.9	920	12	BE785343	BE785343	601478732	C 516	19	0.9	263	9	AA090756	AA090756	603003.0
C 444	19	0.9	923	14	BO887781	BO887781	AGENCOURT	C 517	19	0.9	263	13	BM155124	BM155124	603003.0

518	18	0.8	264	10	BB341316	BB341316	C 591	18	0.8	340	9	A1396637	A1396637	FB15404.x
519	18	0.8	264	14	BQ330257	BQ330257	592	18	0.8	340	17	AQ646105	AQ646105	RPC193-Bc
520	18	0.8	269	9	AL044874	AL044874	593	18	0.8	341	17	AQ008733	AQ008733	RPCC11-22
521	18	0.8	269	13	B1317878	B1317878	C 594	18	0.8	342	12	BG303148	BG303148	FL61F03.x
522	18	0.8	270	9	AA443041	AA443041	595	18	0.8	342	17	AQ848303	AQ848303	LMATFV1.1
523	18	0.8	270	17	AQ070090	AQ070090	596	18	0.8	343	17	AQ848303	AQ848303	LMATFV1.1
524	18	0.8	271	10	AM085416	AM085416	597	18	0.8	344	9	BB852967	BB852967	BB852967
525	18	0.8	274	10	BM194477	BM194477	C 598	18	0.8	346	9	A1332078	A1332078	FA93C01.x
526	18	0.8	278	12	BF137557	BF137557	C 599	18	0.8	346	9	A1332078	A1332078	FA93C01.x
527	18	0.8	278	12	BF137557	BF137557	C 600	18	0.8	347	12	BF511970	BF511970	UI-H-B14-
528	18	0.8	280	10	AV524608	AV524608	C 601	18	0.8	348	9	A1337714	A1337714	NYZ061.Ap
529	18	0.8	280	12	BG588558	BG588558	C 602	18	0.8	349	13	B1318114	B1318114	FGS9F07.x
530	18	0.8	280	12	BF106648	BF106648	C 603	18	0.8	349	13	B172624	B172624	FS02C02.x
531	18	0.8	282	10	AM226648	AM226648	C 604	18	0.8	349	14	BO801082	BO801082	WHE2810.B
532	18	0.8	282	10	AM715652	AM715652	C 605	18	0.8	350	14	BF108111	BF108111	FS42B06.x
533	18	0.8	282	10	BM180176	BM180176	C 606	18	0.8	350	14	BO348857	BO348857	CM4-HT039
534	18	0.8	282	10	BG116556	BG116556	C 607	18	0.8	352	14	D74652	D74652	CEUK083F9F
535	18	0.8	282	12	BF616556	BF616556	C 608	18	0.8	354	10	AW778714	AW778714	FK40G10.x
536	18	0.8	285	12	BF616556	BF616556	C 609	18	0.8	355	10	AW116676	AW116676	FL18E09.x
537	18	0.8	285	14	BQ330258	BQ330258	C 610	18	0.8	355	13	BI982437	BI982437	FE99B06.x
538	18	0.8	286	9	A1313745	A1313745	C 611	18	0.8	357	13	BM081981	BM081981	fu97G02.x
539	18	0.8	286	13	BM157398	BM157398	C 612	18	0.8	357	13	BM081981	BM081981	fu97G02.x
540	18	0.8	287	14	BQ092126	BQ092126	C 613	18	0.8	357	17	AQ007140	AQ007140	CIT-HSP-2
541	18	0.8	287	14	BQ093423	BQ093423	C 614	18	0.8	358	14	BQ092335	BQ092335	Ez40D07.x
542	18	0.8	289	9	A1322230	A1322230	C 615	18	0.8	358	17	AQ044991	AQ044991	RPCC11-33
543	18	0.8	289	9	AA600577	AA600577	C 616	18	0.8	359	17	A2301682	A2301682	Gm UMB001
544	18	0.8	290	10	BB608844	BB608844	C 617	18	0.8	360	9	A1396801	A1396801	FD14C06.x
545	18	0.8	292	9	A1583039	A1583039	C 618	18	0.8	360	13	B1318156	B1318156	FG60N01.x
546	18	0.8	294	10	BB264279	BB264279	C 619	18	0.8	360	17	AQ956947	AQ956947	LERAM94TR
547	18	0.8	294	13	BM533043	BM533043	C 620	18	0.8	361	9	AJ280664	AJ280664	4A3A-AAS-
548	18	0.8	296	10	AM758569	AM758569	C 621	18	0.8	362	13	BI882567	BI882567	FN03D03.x
549	18	0.8	296	14	BM132175	BM132175	C 622	18	0.8	362	13	BI317901	BI317901	FG70F01.x
550	18	0.8	296	14	BM132175	BM132175	C 623	18	0.8	363	17	CNS0331N	CNS0331N	Tetradon
551	18	0.8	298	10	AM936345	AM936345	C 624	18	0.8	364	17	AM567052	AM567052	FK32C07.x
552	18	0.8	300	17	TA97E08P	TA97E08P	C 625	18	0.8	365	10	AM567473	AM567473	FK38H11.x
553	18	0.8	301	10	BM477818	BM477818	C 626	18	0.8	365	10	AM567544	AM567544	FK29A05.x
554	18	0.8	301	10	BM477818	BM477818	C 627	18	0.8	365	10	AM567544	AM567544	FK29A05.x
555	18	0.8	303	13	BM157365	BM157365	C 628	18	0.8	366	10	AM567005	AM567005	FK36F08.x
556	18	0.8	304	13	BM157365	BM157365	C 629	18	0.8	366	10	AM567005	AM567005	FK36F08.x
557	18	0.8	304	13	BM157365	BM157365	C 630	18	0.8	368	10	AM566953	AM566953	FK19D07.x
558	18	0.8	304	13	BM157365	BM157365	C 631	18	0.8	368	12	BF514196	BF514196	UI-H-BW1-
559	18	0.8	305	10	AA736716	AA736716	C 632	18	0.8	368	14	BM778009	BM778009	fy26E06.x
560	18	0.8	309	9	AA736716	AA736716	C 633	18	0.8	369	10	AV525109	AV525109	AV525109
561	18	0.8	310	9	A1583645	A1583645	C 634	18	0.8	369	17	AO127029	AO127029	HS 3041.B
562	18	0.8	310	9	A1583645	A1583645	C 635	18	0.8	370	10	BM203433	BM203433	sf30D11.Y
563	18	0.8	311	9	A1583645	A1583645	C 636	18	0.8	371	13	BI866174	BI866174	fu01C07.x
564	18	0.8	311	9	A1583645	A1583645	C 637	18	0.8	372	9	AA517969	AA517969	vi01H03.x
565	18	0.8	311	13	BM159070	BM159070	C 638	18	0.8	372	10	AM567002	AM567002	FK30F05.x
566	18	0.8	313	9	AL384660	AL384660	C 639	18	0.8	372	14	W66656	W66656	me23d12.x
567	18	0.8	314	12	BF756577	BF756577	C 640	18	0.8	375	10	AM154230	AM154230	fi24b11.x
568	18	0.8	314	12	BF756577	BF756577	C 641	18	0.8	375	10	AM154230	AM154230	fi24b11.x
569	18	0.8	314	13	BM159070	BM159070	C 642	18	0.8	375	10	AM154230	AM154230	fi24b11.x
570	18	0.8	315	13	BM159070	BM159070	C 643	18	0.8	375	12	BM307545	BM307545	fi159F02.x
571	18	0.8	315	13	BM159070	BM159070	C 644	18	0.8	375	12	BM307545	BM307545	fi159F02.x
572	18	0.8	317	9	BM159070	BM159070	C 645	18	0.8	375	12	BM307545	BM307545	fi159F02.x
573	18	0.8	318	10	BM159070	BM159070	C 646	18	0.8	376	12	BM307545	BM307545	fi159F02.x
574	18	0.8	319	12	BM159070	BM159070	C 647	18	0.8	377	12	BM307545	BM307545	fi159F02.x
575	18	0.8	321	9	AA144221	AA144221	C 648	18	0.8	378	9	AA517969	AA517969	vi01H03.x
576	18	0.8	321	13	BM159070	BM159070	C 649	18	0.8	378	14	BM862424	BM862424	mgcm003XG
577	18	0.8	322	9	AA144221	AA144221	C 650	18	0.8	379	9	A1376394	A1376394	lc29e08.x
578	18	0.8	322	10	BM159070	BM159070	C 651	18	0.8	379	12	BF779381	BF779381	h
579	18	0.8	323	10	BM159070	BM159070	C 652	18	0.8	379	13	BM159070	BM159070	h
580	18	0.8	323	10	BM159070	BM159070	C 653	18	0.8	380	13	BM159070	BM159070	h
581	18	0.8	323	10	BM159070	BM159070	C 654	18	0.8	380	13	BM159070	BM159070	h
582	18	0.8	323	10	BM159070	BM159070	C 655	18	0.8	381	10	AM485383	AM485383	65290 MAR
583	18	0.8	323	12	BF569617	BF569617	C 656	18	0.8	382	13	BM159070	BM159070	h
584	18	0.8	323	12	BF569617	BF569617	C 657	18	0.8	384	9	AA605626	AA605626	fa30h10.s
585	18	0.8	323	12	BF569617	BF569617	C 658	18	0.8	384	12	BM159070	BM159070	h
586	18	0.8	323	12	BF569617	BF569617	C 659	18	0.8	386	10	AM154494	AM154494	mac11h12
587	18	0.8	323	12	BF569617	BF569617	C 660	18	0.8	386	13	BM159070	BM159070	h
588	18	0.8	323	12	BF569617	BF569617	C 661	18	0.8	387	9	A1332081	A1332081	fa93C06.x
589	18	0.8	323	12	BF569617	BF569617	C 662	18	0.8	387	17	BM159070	BM159070	h
590	18	0.8	323	12	BF569617	BF569617	C 663	18	0.8	388	9	A1332208	A1332208	NYZ168.Ap

C 664	18	0.8	388	13	BI474394	BI474394 f96d05.x	737	18	0.8	425	17	B86909	B86909 RPCI11-26u1
C 665	18	0.8	388	17	BH018315	BH018315 L142K.d.H	C 738	18	0.8	426	13	BI17924	BI17924 f971c02.x
C 666	18	0.8	390	10	AV591715	AV591715 AV591715	C 739	18	0.8	426	13	BI882960	BI882960 f01d06.x
C 667	18	0.8	390	10	AM886192	AM886192 RC5-OT007	C 740	18	0.8	428	10	AM604922	AM604922 RC0-CT037
C 668	18	0.8	391	9	AI461392	AI461392 f464h03.x	C 741	18	0.8	428	10	AM662374	AM662374 RC0-CT037
C 669	18	0.8	392	9	AJ503218	AJ503218 f464h03.x	C 742	18	0.8	428	14	BQ515564	BQ515564 EST622979
C 670	18	0.8	392	12	BG798218	BG798218 f07b01.y	C 743	18	0.8	429	13	BI433161	BI433161 EST359922
C 671	18	0.8	392	14	RO8011	RO8011 yf17a07.st	C 744	18	0.8	429	14	BQ515563	BQ515563 EST622978
C 672	18	0.8	393	9	AI330884	AI330884 f0d5c11.x	C 745	18	0.8	429	17	BH223514	BH223514 106611300
C 673	18	0.8	393	9	AI331376	AI331376 f468a11.x	C 746	18	0.8	429	17	FR0044532	FR0044532 fugu.rubr
C 674	18	0.8	393	9	AI385060	AI385060 f0d6f05.x	C 747	18	0.8	430	9	AI741515	AI741515 w921c01.x
C 675	18	0.8	393	10	AA441862	AA441862 zw62d02.r	C 748	18	0.8	430	9	AI943751	AI943751 618045A12
C 676	18	0.8	393	10	AM175461	AM175461 f125d09.x	C 749	18	0.8	430	13	EM028966	EM028966 f98h03.r
C 677	18	0.8	393	12	BG882969	BG882969 f32b06.x	C 750	18	0.8	432	9	AA717831	AA717831 v98h03.r
C 678	18	0.8	393	14	BM797887	BM797887 K-EST0081	C 751	18	0.8	432	13	BI562946	BI562946 f6q7h09.x
C 679	18	0.8	393	14	BQ093716	BQ093716 f45h08.x	C 752	18	0.8	432	13	BI864260	BI864260 f19g01.x
C 680	18	0.8	393	17	AZ304852	AZ304852 1M0005F19	C 753	18	0.8	432	13	BI864260	BI864260 f19g01.x
C 681	18	0.8	394	9	AA954411	AA954411 on85a08.b	C 754	18	0.8	433	13	BI841415	BI841415 f955g02.x
C 682	18	0.8	394	9	AI333065	AI333065 f493a11.x	C 755	18	0.8	433	14	BM856200	BM856200 K-EST0139
C 683	18	0.8	394	9	AI882689	AI882689 f015b10.y	C 756	18	0.8	433	14	BQ393795	BQ393795 NISC.ng05
C 684	18	0.8	394	10	AM420437	AM420437 f181e12.x	C 757	18	0.8	434	14	R82218	R82218 y17a01.r1
C 685	18	0.8	394	10	AM420437	AM420437 f181e12.x	C 758	18	0.8	437	12	B8223837	B8223837 NCST3a27
C 686	18	0.8	394	14	BQ616909	BQ616909 f4b01f12.x	C 759	18	0.8	438	10	AM677615	AM677615 DG1.8.B10
C 687	18	0.8	395	9	AI396621	AI396621 f015b10.x	C 760	18	0.8	438	13	BI864373	BI864373 f195b06.x
C 688	18	0.8	395	10	BES24893	BES24893 M55F3STM	C 761	18	0.8	438	14	BQ167623	BQ167623 WHE0081.C
C 689	18	0.8	395	17	BH071136	BH071136 RPCI-24-2	C 762	18	0.8	439	10	AM525853	AM525853 UI-R-BJ09
C 690	18	0.8	395	17	AQ409333	AQ409333 HS.5067.A	C 763	18	0.8	441	10	AM677571	AM677571 DG1.8.H09
C 691	18	0.8	396	9	AI331263	AI331263 f487b05.x	C 764	18	0.8	441	17	AZ049683	AZ049683 GSTCT072
C 692	18	0.8	396	10	BES57308	BES57308 P11-71.H0	C 765	18	0.8	442	10	B8600291	B8600291 P11.94.F0
C 693	18	0.8	396	17	AQ009088	AQ009088 RPCI11-22	C 766	18	0.8	442	13	BI294608	BI294608 UI-R-DK0
C 694	18	0.8	397	12	BE753153	BE753153 206068.MA	C 767	18	0.8	442	17	B22546	B22546 F2G05TF.IGF
C 695	18	0.8	397	13	BI841412	BI841412 f455f10.x	C 768	18	0.8	443	10	AM975360	AM975360 EST87468
C 696	18	0.8	397	13	BI841412	BI841412 f455f10.x	C 769	18	0.8	443	9	BM386075	BM386075 UI-R-DN1-
C 697	18	0.8	397	13	BI843467	BI843467 f490a06.x	C 770	18	0.8	444	9	AA545284	AA545284 y194a12.b
C 698	18	0.8	398	9	AI330336	AI330336 f481a11.x	C 771	18	0.8	444	14	W57704	W57704 z018b02.st
C 699	18	0.8	398	9	AI330869	AI330869 f0d5b05.x	C 772	18	0.8	445	10	AM886190	AM886190 RC5-OT007
C 700	18	0.8	398	10	AM778715	AM778715 f480g11.x	C 773	18	0.8	445	13	BI707672	BI707672 f36b08.x
C 701	18	0.8	399	10	AM886196	AM886196 RC5-OT007	C 774	18	0.8	445	14	W98555	W98555 mg14d02.r1
C 702	18	0.8	399	9	AI385082	AI385082 f0d6h09.x	C 775	18	0.8	445	17	AZ352576	AZ352576 1M0091A09
C 703	18	0.8	400	9	AI330522	AI330522 f482g07.x	C 776	18	0.8	446	9	AI506751	AI506751 vms6e05.x
C 704	18	0.8	401	9	AI331795	AI331795 f01e04.x	C 777	18	0.8	447	9	AI752740	AI752740 c18b602.x
C 705	18	0.8	401	9	AA101998	AA101998 z178a06.r	C 778	18	0.8	447	12	B8972219	B8972219 f015b10.f
C 706	18	0.8	402	10	AM778710	AM778710 f480g05.x	C 779	18	0.8	447	13	BM081796	BM081796 f094h01.x
C 707	18	0.8	404	13	BM081889	BM081889 f096c02.x	C 780	18	0.8	447	17	B84170	B84170 RPCI11-22G7
C 708	18	0.8	404	17	AZ657923	AZ657923 1M0534K24	C 781	18	0.8	448	9	AA606120	AA606120 f416a07.b
C 709	18	0.8	405	13	BI429882	BI429882 f473d09.y	C 782	18	0.8	448	13	BI702078	BI702078 s8g91909.
C 710	18	0.8	405	13	BI429882	BI429882 f473d09.y	C 783	18	0.8	448	17	AQ504349	AQ504349 RPCI-11-2
C 711	18	0.8	405	14	H86709	H86709 yf08a10.r1	C 784	18	0.8	449	10	AM669419	AM669419 112783.MA
C 712	18	0.8	406	12	BF559749	BF559749 UI-R-E1.G	C 785	18	0.8	449	14	BQ576977	BQ576977 f195f01.x
C 713	18	0.8	406	13	BI429890	BI429890 f476e09.x	C 786	18	0.8	451	13	BI474459	BI474459 f195f01.x
C 714	18	0.8	407	9	AU174950	AU174950 f476e09.x	C 787	18	0.8	452	10	BE363310	BE363310 WSI.61.A0
C 715	18	0.8	408	17	AQ787388	AQ787388 HS.3224.B	C 788	18	0.8	452	13	BI864477	BI864477 f195f01.x
C 716	18	0.8	409	9	AI415823	AI415823 f01e05.x	C 789	18	0.8	453	13	BI841550	BI841550 f195f01.x
C 717	18	0.8	411	12	BE991893	BE991893 UI-M-BZ1-	C 790	18	0.8	453	13	BI843585	BI843585 f195f01.x
C 718	18	0.8	411	13	BM289291	BM289291 S30866.MA	C 791	18	0.8	453	13	BI864445	BI864445 f195f01.x
C 719	18	0.8	412	14	BM662423	BM662423 mgcm003xG	C 792	18	0.8	453	17	AO816232	AO816232 HS.5443.B
C 720	18	0.8	414	14	C73753	C73753 C73753.Rice	C 793	18	0.8	454	9	AA605954	AA605954 f41p04.3
C 721	18	0.8	415	14	BQ451845	BQ451845 PESTOADO	C 794	18	0.8	454	10	AM940655	AM940655 CH17344.3
C 722	18	0.8	417	12	BG892532	BG892532 f474f07.x	C 795	18	0.8	454	10	BB861395	BB861395 BB861395
C 723	18	0.8	417	13	BF261616	BF261616 u248d08.x	C 796	18	0.8	454	13	BI429930	BI429930 f477c12.x
C 724	18	0.8	417	13	BI275463	BI275463 UI-R-CX0-	C 797	18	0.8	454	13	BI474409	BI474409 f195f01.x
C 725	18	0.8	418	10	BE366814	BE366814 P11.41.G0	C 798	18	0.8	454	13	BI707719	BI707719 f195f01.x
C 726	18	0.8	418	12	BG892651	BG892651 f477f08.x	C 799	18	0.8	454	13	BI707985	BI707985 f477f08.x
C 727	18	0.8	418	13	BI334245	BI334245 602997754	C 800	18	0.8	454	13	BI784634	BI784634 g4f57d06.
C 728	18	0.8	419	10	AM710611	AM710611 esa01ne.f	C 801	18	0.8	454	13	BI844015	BI844015 f477f08.x
C 729	18	0.8	419	13	BI189514	BI189514 f477f08.x	C 802	18	0.8	454	13	BI982460	BI982460 f477f08.x
C 730	18	0.8	419	13	BI189514	BI189514 f477f08.x	C 803	18	0.8	454	14	T97908	T97908 y654c11.st
C 731	18	0.8	419	13	BI479399	BI479399 f477f08.x	C 804	18	0.8	455	9	AI444468	AI444468 f477f08.x
C 732	18	0.8	422	9	AI288259	AI288259 f180c09.x	C 805	18	0.8	455	13	BI708187	BI708187 f477f08.x
C 733	18	0.8	422	9	AI397482	AI397482 f180c09.x	C 806	18	0.8	455	13	BI843989	BI843989 f477f08.x
C 734	18	0.8	422	10	AM614916	AM614916 nh70f01.x	C 807	18	0.8	455	13	BI864500	BI864500 f477f08.x
C 735	18	0.8	422	14	R82176	R82176 yj17a01.st	C 808	18	0.8	455	13	BI866074	BI866074 f477f08.x
C 736	18	0.8	423	14	BM671510	BM671510 UI-E-CX1-	C 809	18	0.8	455	13	BI190987	BI190987 f477f08.x





Db 704 GGCAGTGGAAAGTGAACGAGGTGGTTACAAAGTTTAAATGAAAGATTAATACAGGCTCT 763  
 QY 667 GGAAGAATTTCTTGAAGTCAGAGCAAGAGAGAGAAAGTAGTAGTACTCAAGACCA 726  
 Db 764 GGAAGAATTTCTTGAAGTCAGAGCAAGAGAGAGAAAGTAGTAGTACTCAAGACCA 823  
 QY 727 AAAGTACCTACATACCCCTCTCTCCAGCTGAGAGATGAGAGCTCAATCTTTGACATTTAT 786  
 Db 824 AAAGTACCTACATACCCCTCTCTCCAGCTGAGAGATGAGAGCTCAATCTTTGACATTTAT 883  
 QY 787 CAGACAGGATTAATCTTGAAGCAATATGAGCACTATTTCTTGAAGAGTCTGAGCAATGAT 846  
 Db 884 CAGACAGGATTAATCTTGAAGCAATATGAGCACTATTTCTTGAAGAGTCTGAGCAATGAT 943  
 QY 847 GCACACACGACATTTCTGACTTTTGAAGAGTATCTCTGACAGACTGAAATTAACAAC 906  
 Db 944 GCACACACGACATTTCTGACTTTTGAAGAGTATCTCTGACAGACTGAAATTAACAAC 1003  
 QY 907 ATTGCTAAAGCTGGTTATATCTAAGCTTACTCTCTGTGCAAAAATACAGTATTTCTATCATA 966  
 Db 1004 ATTGCTAAAGCTGGTTATATCTAAGCTTACTCTCTGTGCAAAAATACAGTATTTCTATCATA 1063  
 QY 967 CTTCGAGAGAGATTTGATGAGTCTGTCTCAAAACAGGGTCTGGAAGAGCTGGGCTTTT 1026  
 Db 1064 CTTCGAGAGAGATTTGATGAGTCTGTCTCAAAACAGGGTCTGGAAGAGCTGGGCTTTT 1123  
 QY 1027 CTCTCTCAATTTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086  
 Db 1124 CTCTCTCAATTTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1183  
 QY 1087 TTGACAGAAACAGAGTGTATATTTGTAGACACCAATCTGAGAAATTTGTCAACAGATTTAT 1146  
 Db 1184 TTGACAGAAACAGAGTGTATATTTGTAGACACCAATCTGAGAAATTTGTCAACAGATTTAT 1243  
 QY 1147 TTGGAAGCCAGAAATTTTCTTTTGGAGCTTTGTGAAGAGCTTTGTATATATGGGGGA 1206  
 Db 1244 TTGGAAGCCAGAAATTTTCTTTTGGAGCTTTGTGAAGAGCTTTGTATATATGGGGGA 1303  
 QY 1207 ACCAGCTGGGACATTTCAATTTGCAAAATAGTACAAAGGCTGTAAATATATATGAGTACT 1266  
 Db 1304 ACCAGCTGGGACATTTCAATTTGCAAAATAGTACAAAGGCTGTAAATATATATGAGTACT 1363  
 QY 1267 CCTGGAAGACTGATGATATCATATAGGCAAGAAAGATTTGTCTCAAAACAGATCAATAAC 1326  
 Db 1364 CC-GGAAGACTGATGATATCATATAGGCAAGAAAGATTTGTCTCAAAACAGATCAATAAC 1422  
 QY 1327 TTATGTTGGATGAAGCTGATCGCATGTTGGATGAGTGGTGGTGGTGGTGGTGGTGGTGGT 1386  
 Db 1423 TTATGTTGGATGAAGCTGATCGCATGTTGGATGAGTGGTGGTGGTGGTGGTGGTGGTGGT 1482  
 QY 1387 TTATGTTGGTGGCCAGGAATGCAATCAAGAAACAGGCCAACCCCTTATGTTCAAGTCA 1446  
 Db 1483 TTATGTTGGTGGCCAGGAATGCAATCAAGAAACAGGCCAACCCCTTATGTTCAAGTCA 1542  
 QY 1447 ACTTTTCCAGAGGAAATTTCAAGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1506  
 Db 1543 ACTTTTCCAGAGGAAATTTCAAGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1602  
 QY 1507 GTTGTCTGTGGACAAGTGGGTGAGCATGTAGAGATTTGACAGACCGTTCTTCCAAATT 1566  
 Db 1603 GTTGTCTGTGGACAAGTGGGTGAGCATGTAGAGATTTGACAGACCGTTCTTCCAAATT 1662  
 QY 1567 GGGCAGTCTTCAAAAAGAGAAAGCTCGTTGAATTTCTGGAACATATGAGGATGAAAGA 1626  
 Db 1663 GGGCAGTCTTCAAAAAGAGAAAGCTCGTTGAATTTCTGGAACATATGAGGATGAAAGA 1722  
 QY 1627 ACTATGCTCTTGTGGAACATTAAGAAAAAGCAGATTTTACTGCACTTTCTTTGTCAA 1686  
 Db 1723 ACTATGCTCTTGTGGAACATTAAGAAAAAGCAGATTTTACTGCACTTTCTTTGTCAA 1782  
 QY 1687 GAAAAAATATCAACTACAGTATTCATGTGATCGGAGACAGAGAGCGGAGCAAGCT 1746  
 Db 1783 GAAAAAATATCAACTACAGTATTCATGTGATCGGAGACAGAGAGCGGAGCAAGCT 1842

QY 1747 CTTCGAGATTTTTCGCTTTGGAAGAGTCCAGTCTTGTGCTACTTCACTAGCTGCCAGA 1806  
 Db 1843 CTTCGAGATTTTTCGCTTTGGAAGAGTCCAGTCTTGTGCTACTTCACTAGCTGCCAGA 1902  
 QY 1807 GGGCTGATATTTGAAATTTGCAACATGTTATCATTTTGTGATCTTCTTACCATTTGAT 1866  
 Db 1903 GGGCTGATATTTGAAATTTGCAACATGTTATCATTTTGTGATCTTCTTACCATTTGAT 1962  
 QY 1867 GAATATGTTATTCGAAATTTGGGGGTACTGGTCTGTGGGAAATATCTGACAGCAATTTCC 1926  
 Db 1963 GAATATGTTATTCGAAATTTGGGGGTACTGGTCTGTGGGAAATATCTGACAGCAATTTCC 2022  
 QY 1927 TTTTGTATCTTGAATTCGATTAACCATTTTGAACAAGCTCTAGTAAAGATTTGACAGAT 1986  
 Db 2023 TTTTGTATCTTGAATTCGATTAACCATTTTGAACAAGCTCTAGTAAAGATTTGACAGAT 2082  
 QY 1987 GCTCAACAGAGATGTTCTTCATGATGTTGAAGAAATTCCTTTAGTACATACATTCCTGGC 2046  
 Db 2083 GCTCAACAGAGATGTTCTTCATGATGTTGAAGAAATTCCTTTAGTACATACATTCCTGGC 2142  
 QY 2047 TTCAGTGTAGTACAGAGAAACGTTTGCATGCTGATACAGAAAGGCAAGAGC 2106  
 Db 2143 TTCAGTGTAGTACAGAGAAACGTTTGCATGCTGATACAGAAAGGCAAGAGC 2202  
 QY 2107 ACTTTGACACAGCTGGTCTTCTTCTTCTGAC 2137  
 Db 2203 ACTTTGACACAGCTGGTCTTCTTCTTCTTCTGAC 2233

RESULT 2  
 BG718234  
 LOCUS  
 DEFINITION  
 mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 682)  
 NIH-MGC http://img.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
 Toshitaki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 http://image.jnl.gov  
 Plate: LHAM10746 row: a column: 03  
 High quality sequence stop: 682.  
 Location/Qualifiers  
 1. 682  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="IMAGE:4828298"  
 /clone\_1b="NIH-MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescriptR (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
 ); Oligo-dt primed using primer 5'-TTTCTTTTCTTTTCTTN-3',  
 size-selected for average insert size 2.2 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NHGRI/NHGR1, National

## FEATURES

source

BASE COUNT 205 a 109 c 204 g 164 t  
 INSTITUTEs of Health). Note: this is a NIH\_MGC Library."

Query Match 28.4%; Score 616; DB 13; Length 693;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 GATGAATTTGGGAGAGAGAAATCAACCCATATATGCTTCCTCATATGTTCCCATATTTGAG 66  
 67 GATGAATTTGGGAGAGAGAAATCAACCCATATATGCTTCCTCATATGTTCCCATATTTGAG 126  
 67 AAGGATAGTATTTCTGGAGAAATGAGACATTTTAAAGAGACTCCAGCTTCATCATCA 126  
 127 AAGGATAGTATTTCTGGAGAAATGAGACATTTTAAAGAGACTCCAGCTTCATCATCA 166  
 127 GAAATGATATGACCTTCTCGAAGAGATCATTTCAATGAAAGTGGATTGGCTCTGGG 186  
 187 GAAATGATATGACCTTCTCGAAGAGATCATTTCAATGAAAGTGGATTGGCTCTGGG 246  
 187 CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATATACATCCAGATG 246  
 247 CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATATACATCCAGATG 306  
 247 GGTGCTTTGGAGTTGGAAAGAGTTTGGAAACAGAGTTTTCGAACAGAGTTTGA 306  
 307 GGTGCTTTGGAGTTGGAAAGAGTTTGGAAACAGAGTTTTCGAACAGAGTTTGA 366  
 307 GATGATATAGCTCTGCTGTTCTGGAGAGATCTAGTATATAGCTGCGAAGATATCCACA 366  
 367 GATGATATAGCTCTGCTGTTCTGGAGAGATCTAGTATATAGCTGCGAAGATATCCACA 426  
 367 CGGAACAGAGGTTTCCAGAGAGGCGGCTATGAGATGGAATTAATTCAGAGCTTCA 426  
 427 CGGAACAGAGGTTTCCAGAGAGGCGGCTATGAGATGGAATTAATTCAGAGCTTCA 486  
 427 GGGCATATACAGAGAGGTTGAGAGAGTATTTCCGAGTTGCCGTGAGAGATTGGCTCA 546  
 487 GGGCATATACAGAGAGGTTGAGAGAGTATTTCCGAGTTGCCGTGAGAGATTGGCTCA 546  
 487 GGAAGTCCAAATATATGATGACCCAGAGCAATGATGACGCGCATGCGGCTTTT 546  
 547 GGAAGTCCAAATATATGATGACCCAGAGCAATGATGACGCGCATGCGGCTTTT 606  
 547 GGTCTAGAGAGACAGATTAATGAGTGAAGTATAGTGAATTAATTCAGAGAGT 606  
 607 GGTCTAGAGAGACAGATTAATGAGTGAAGTATAGTGAATTAATTCAGAGAGT 666  
 607 GGCAGTGAAGTGAAC 622  
 667 GGCAGTGAAGTGAAC 682

RESULT 3  
 LOCUS B1560849 693 bp mRNA linear EST 05-SEP-2001  
 DEFINITION 603254018F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5256444 5',  
 mRNA sequence.  
 ACCESSION B1560849  
 VERSION B1560849.1 GI:15448163  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 693)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-rc@mail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM11750 row: c column: 05  
 High quality sequence stop: 691.  
 Location/Qualifiers  
 1. 693  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5256444"  
 /clone\_lib="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescript (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (Gcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
 size-selected for average insert size 2.2 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 207 a 111 c 207 g 168 t

Query Match 28.4%; Score 616; DB 13; Length 693;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 GATGAATTTGGGAGAGAGAAATCAACCCATATATGCTTCCTCATATGTTCCCATATTTGAG 66  
 78 GATGAATTTGGGAGAGAGAAATCAACCCATATATGCTTCCTCATATGTTCCCATATTTGAG 137  
 67 AAGGATAGTATTTCTGGAGAAATGAGACATTTTAAAGAGACTCCAGCTTCATCATCA 126  
 138 AAGGATAGTATTTCTGGAGAAATGAGACATTTTAAAGAGACTCCAGCTTCATCATCA 197  
 127 GAAATGATATGACCTTCTCGAAGAGATCTAGTATATAGCTGCGAAGATATCCACA 186  
 198 GAAATGATATGACCTTCTCGAAGAGATCTAGTATATAGCTGCGAAGATATCCACA 257  
 187 CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATATACATCCAGATG 246  
 258 CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATATACATCCAGATG 317  
 247 GGTGCTTTGGAGTTGGAAAGAGTTTGGAAACAGAGTTTTCGAACAGAGTTTGA 306  
 318 GGTGCTTTGGAGTTGGAAAGAGTTTGGAAACAGAGTTTTCGAACAGAGTTTGA 377  
 307 GATGATATAGCTCTGCTGTTCTGGAGAGATCTAGTATATAGCTGCGAAGATATCCACA 366  
 378 GATGATATAGCTCTGCTGTTCTGGAGAGATCTAGTATATAGCTGCGAAGATATCCACA 437  
 367 CGGAACAGAGGTTTCCAGAGAGGCGGCTATGAGATGGAATTAATTCAGAGCTTCA 426  
 438 CGGAACAGAGGTTTCCAGAGAGGCGGCTATGAGATGGAATTAATTCAGAGCTTCA 497  
 427 GGGCATATACAGAGAGTGAAGAGTATTTCCGAGTTGCCGTGAGAGATTGGCTCA 486  
 498 GGGCATATACAGAGAGTGAAGAGTATTTCCGAGTTGCCGTGAGAGATTGGCTCA 557  
 487 GGAAGTCCAAATATATGATGACCCAGAGCAATGATGACGCGCATGCGGCTTTT 546  
 558 GGAAGTCCAAATATATGATGACCCAGAGCAATGATGACGCGCATGCGGCTTTT 617  
 547 GGTCTAGAGAGACAGATTAATGAGTGAAGTATAGTGAATTAATTCAGAGAGT 606  
 618 GGTCTAGAGAGACAGATTAATGAGTGAAGTATAGTGAATTAATTCAGAGAGT 677  
 607 GGCAGTGAAGTGAAC 622



Db 678 GGCAGTGGAAATGAAC 693

RESULT 4  
LOCUS BG717110 798 bp mRNA linear EST 08-MAY-2001  
DEFINITION 602689191F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4821503 5',  
mRNA sequence.  
ACCESSION BG717110  
VERSION BG717110.1 GI:13996297  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 798)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM410728 row: e column: 24  
High quality sequence stop: 770.  
Location/Qualifiers  
1..798  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4821503"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptPR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI; XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVA-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 254 a 124 c 233 g 186 t 1 others  
ORIGIN

Query Match 28.2%; Score 612; DB 12; Length 798;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 662; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 TGGGAAGCAGAAATCAACCTCATATGCTTCTTATTTCCATATTTGGAAGATAGG 75  
DB 66 TGGGAAGCAGAAATCAACCTCATATGCTTCTTATTTCCATATTTGGAAGATAGG 125  
QY 76 TATTTGGAGAAATGAGACAAATTTTACAGAGCTCCAGCTTCATCATCAGAAATGAT 135  
DB 126 TATTTGGAGAAATGAGACAAATTTTACAGAGCTCCAGCTTCATCATCAGAAATGAT 185  
QY 136 GATGACCTTCTGAGAGAGATCATTTTCAGAAAGGATTTGCTGGGCGGAATTTT 195  
DB 186 GATGACCTTCTGAGAGAGATCATTTTCAGAAAGGATTTGCTGGGCGGAATTTT 245  
QY 196 GGAAGACAGAGATCTGCTGAGTGTATTAACGAGATATATCATCAATGGGTGTTTT 255  
DB 246 GAAAGACAGAGATCTGCTGAGTGTATTAACGAGATATATCATCAATGGGTGTTTT 305  
QY 256 GGAAGTGGAAAGGTTTTGGAACAGAGGTTTTTCAACGCGAGGTTTGAAGTGTGAT 315  
DB 306 GGAAGTGGAAAGGTTTTGGAACAGAGGTTTTTCAACGCGAGGTTTGAAGTGTGAT 365

QY 316 AGCTGTGTTTTCTGAGAGAGTCTAGTAATGACTGCGAAGATATCAACGGAACAGA 375  
DB 366 AGCTGTGTTTTCTGAGAGAGTCTAGTAATGACTGCGAAGATATCAACGGAACAGA 425  
QY 376 GGGTTTTCCAGAGAGCGCGCTATGAGATGGAATAATTCAGAGCTTCAGGCCATAC 435  
DB 426 GGGTTTTCCAGAGAGCGCGCTATGAGATGGAATAATTCAGAGCTTCAGGCCATAC 485  
QY 436 AGAAGAGGTGAAGAGGTAGTTCCGAGGTCCCGGAGAGATTTGGCTAGGAAGTCA 495  
DB 486 AGAAGAGGTGAAGAGGTAGTTCCGAGGTCCCGGAGAGATTTGGCTAGGAAGTCA 545  
QY 496 AATAATGACTTACCCACAGCAATGTATGACGCGACTGTGGCTTTTGGTTTAGA 555  
DB 546 AATAATGACTTACCCACAGCAATGTATGACGCGACTGTGGCTTTTGGTTTAGA 605  
QY 556 AGACGAGTATTAAGTGGACAGGTAAATGATCTTCAAGCAGAAAGTGGAGTGA 615  
DB 606 AGACGAGTATTAAGTGGACAGGTAAATGATCTTCAAGCAGAAAGTGGAGTGA 665  
QY 616 AGTGAACGAGGTGTTTACAAAGTTAATGAAGATTAACAGCTTCGAAAGAT 675  
DB 666 AGTGAACGAGGTGTTTACAAAGTTAATGAAGATTAACAGCTTCGAAAGAT 725  
QY 676 TCT 678  
DB 726 TCT 728

RESULT 5  
LOCUS BQ429126 792 bp mRNA linear EST 24-MAY-2002  
DEFINITION AGENCOURT\_7901787 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6104663  
5', mRNA sequence.  
ACCESSION BQ429126  
VERSION BQ429126.1 GI:21168202  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 792)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM2342 row: n column: 24  
High quality sequence stop: 496.  
Location/Qualifiers  
1..792  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6104663"  
/clone\_lib="NIH\_MGC\_82"  
/lab\_host="DH10B (TI phase-resistant)"  
/note="Organ: testis; Vector: pNR-LIB (Clontech); Site\_1:  
SfiI (ggccgctcgcc); Site\_2: SfiI (ggccatcggcc); 5' and  
3' adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CACGGCATTATGACC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCGCGCGCGCAATG-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size  
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained  
inserts by PCR. This library was enriched for full-length  
clones and was constructed by Clontech Laboratories (Palo

BASE COUNT 241 a 128 c 242 g 180 t 1 others  
ORIGIN

Query Match 24.6%; Score 535; DB 14; Length 792;  
Best Local Similarity 100.0%; Pred. No. 2.3e-281;  
Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 GATGAAGATTGGGAGAGCAAAATCAACCCATATGCTTCCATATGTTCCCATATTTGAG 66  
Db GATGAAGATTGGGAGAGCAAAATCAACCCATATGCTTCCATATGTTCCCATATTTGAG 126  
67 GATGAAGATTGGGAGAGCAAAATCAACCCATATGCTTCCATATGTTCCCATATTTGAG 126  
QY AAGGATAGTATCTCGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTATCATCA 126  
Db AAGGATAGTATCTCGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTATCATCA 186  
127 GAAATGATGATGACCTTCTCGAAGAGATCATTTCAATGAAAGTGGATTGGCTCTGGG 186  
Db GAAATGATGATGACCTTCTCGAAGAGATCATTTCAATGAAAGTGGATTGGCTCTGGG 246  
187 CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATTAATACATCCAAATG 246  
QY CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATTAATACATCCAAATG 246  
Db CGGAATTTGGAAACAGAGATGCTGTGAGTATATAGGAGATTAATACATCCAAATG 306  
247 GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGCTTTTCAACAGAGTTTGA 306  
QY GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGCTTTTCAACAGAGTTTGA 306  
Db GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGCTTTTCAACAGAGTTTGA 366  
307 GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 366  
QY GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 366  
Db GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 426  
367 CGGAAGCAGAGGTTTTCAGAGAGGCGGCTATGAGATGGAATTAATTAAGAGCTTCA 426  
QY CGGAAGCAGAGGTTTTCAGAGAGGCGGCTATGAGATGGAATTAATTAAGAGCTTCA 426  
Db CGGAAGCAGAGGTTTTCAGAGAGGCGGCTATGAGATGGAATTAATTAAGAGCTTCA 486  
427 GGGCCATACAGAGAGGTTGAGAGAGTATTTCCGAGCTTCCGTTGAGAGATTGCTCTA 486  
QY GGGCCATACAGAGAGGTTGAGAGAGTATTTCCGAGCTTCCGTTGAGAGATTGCTCTA 486  
Db GGGCCATACAGAGAGGTTGAGAGAGTATTTCCGAGCTTCCGTTGAGAGATTGCTCTA 546  
487 GGAAGTCCAAATATATGACTTACAGCCAGAGATGATGACGCCACTGTGGCC 601  
QY GGAAGTCCAAATATATGACTTACAGCCAGAGATGATGACGCCACTGTGGCC 601  
Db GGAAGTCCAAATATATGACTTACAGCCAGAGATGATGACGCCACTGTGGCC 601

RESULT 6  
LOCUS BG719449 747 bp mRNA linear EST 08-MAY-2001  
DEFINITION 602690172F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4822528 5',  
mRNA sequence.  
BG719449  
ACCESSION BG719449.1 GI:13998636  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 747)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LHAM0730 row: p column: 17  
High quality sequence stop: 734.

FEATURES  
source

Location/Qualifiers  
1..747  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4822528"  
/clone\_1b="NIH\_MGC\_97"  
/ab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',  
size-selected for average insert size 2.2 kb and  
normalized to 10^5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 220 a 121 c 226 g 180 t  
ORIGIN

Query Match 23.4%; Score 509; DB 12; Length 747;  
Best Local Similarity 99.8%; Pred. No. 4.1e-267;  
Matches 623; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

7 GATGAAGATTGGGAGAGCAAAATCAACCCATATGCTTCCATATGTTCCCATATTTGAG 66  
Db GATGAAGATTGGGAGAGCAAAATCAACCCATATGCTTCCATATGTTCCCATATTTGAG 162  
103 GATGAAGATTGGGAGAGCAAAATCAACCCATATGCTTCCATATGTTCCCATATTTGAG 162  
QY AAGGATAGTATCTCGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTATCATCA 126  
Db AAGGATAGTATCTCGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTATCATCA 222  
163 AAGGATAGTATCTCGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTATCATCA 222  
QY GAAATGATGATGACCTTCTCGAAGAGATCATTTTCAATGAAAGTGGCTCTGGG 186  
Db GAAATGATGATGACCTTCTCGAAGAGATCATTTTCAATGAAAGTGGCTCTGGG 282  
223 GAAATGATGATGACCTTCTCGAAGAGATCATTTTCAATGAAAGTGGCTCTGGG 282  
QY CGGAATTTTGGAAACAGAGATGCTGTGAGTATATAGGAGATTAATCAATG 246  
Db CGGAATTTTGGAAACAGAGATGCTGTGAGTATATAGGAGATTAATCAATG 342  
283 CGGAATTTTGGAAACAGAGATGCTGTGAGTATATAGGAGATTAATCAATG 342  
QY GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGGTTTTCAGAGCTTCCGTTGAGAGATTGCTCTA 306  
Db GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGGTTTTCAGAGCTTCCGTTGAGAGATTGCTCTA 402  
247 GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGGTTTTCAGAGCTTCCGTTGAGAGATTGCTCTA 402  
QY GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGGTTTTCAGAGCTTCCGTTGAGAGATTGCTCTA 462  
Db GGTGTTTGGAGTTGGAAGAGTTTGAACAGAGGTTTTCAGAGCTTCCGTTGAGAGATTGCTCTA 462  
307 GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 366  
QY GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 366  
Db GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 462  
403 GATGATATAGCTCTGCTTCTGAGAGAGTCTATGATATGATCTGCGAATTAATCAACA 462  
QY CGGAACAGAGGTTTTCAGAGAGGCGGCTATGAGATGGAATTAATTAAGAGCTTCA 425  
Db CGGAACAGAGGTTTTCAGAGAGGCGGCTATGAGATGGAATTAATTAAGAGCTTCA 522  
463 CGGAACAGAGGTTTTCAGAGAGGCGGCTATGAGATGGAATTAATTAAGAGCTTCA 522  
QY AGGGCCATACAGAGAGGTTGAGAGAGTATTTCCGAGCTTCCGTTGAGAGATTGCTCT 485  
Db AGGGCCATACAGAGAGGTTGAGAGAGTATTTCCGAGCTTCCGTTGAGAGATTGCTCT 582  
523 AGGGCCATACAGAGAGGTTGAGAGAGTATTTCCGAGCTTCCGTTGAGAGATTGCTCT 582  
QY AGGAAGTCCAAATATATGACTTACAGCCAGAGATGATGACGCCACTGTGGCC 545  
Db AGGAAGTCCAAATATATGACTTACAGCCAGAGATGATGACGCCACTGTGGCC 642  
583 AGGAAGTCCAAATATATGACTTACAGCCAGAGATGATGACGCCACTGTGGCC 642  
QY TGGTTTCTAGAGACAGATTAAGTGCGACAGGTAATGATCTTCAAGAGCAAG 605  
Db TGGTTTCTAGAGACAGATTAAGTGCGACAGGTAATGATCTTCAAGAGCAAG 702  
643 TGGTTTCTAGAGACAGATTAAGTGCGACAGGTAATGATCTTCAAGAGCAAG 702  
QY TGGCACTGGAAGTGAACGAGGTGCTTCA 635  
Db TGGCACTGGAAGTGAACGAGGTGCTTCA 732  
703 TGGCACTGGAAGTGAACGAGGTGCTTCA 732

RESULT 7  
LOCUS BG717812 657 bp mRNA linear EST 08-MAY-2001  
DEFINITION 602693829F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4826190 5',  
mRNA sequence.

ACCESSION EG717812  
VERSION EG717812.1 GI:13996999  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 657)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM10740 row: 1 column: 07  
High quality sequence stop: 655.  
Location/Qualifiers  
1. 657  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1ib="IMAGE:482149"  
/clone\_1ib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgcag  
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTN-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 200 a 105 c 192 g 160 t  
ORIGIN  
Query Match 22.1%; Score 480; DB 12; Length 657;  
Best Local Similarity 99.8%; Pred. No. 3.2e-251;  
Matches 600; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 7 GATGAAGATTGGAGACCAAAATCAACCTCATATGCTCTTATGTTCCCATATTGAG 66  
DB 57 GATGAAGATTGGAGACCAAAATCAACCTCATATGCTCTTATGTTCCCATATTGAG 116  
QY 67 AAGATAGGATATCTGAGAAATGAGACCAATTTAAGAGAGCTCAGCTTCATCATCA 126  
DB 117 AAGATAGGATATCTGAGAAATGAGACCAATTTAAGAGAGCTCAGCTTCATCATCA 176  
QY 127 GAAATGATGATGACCTCTCGAAGAGATCATATTCATGAAAGTGATTTGCTCTGAG 186  
DB 177 GAAATGATGATGACCTCTCGAAGAGATCATATTCATGAAAGTGATTTGCTCTGAG 236  
QY 187 CGGAATTTTGGAAACGAGATGCTGTGATGATTAAGAGATATATCATCACAAAG 246  
DB 237 CGGAATTTTGGAAACGAGATGCTGTGATGATTAAGAGATATATCATCACAAAG 296  
QY 247 GGTGTTTGGAGTTGAAAGAGTTTGGAAACAGAGTTTCAACAGCAGAGTTGAA 306  
DB 297 GGTGTTTGGAGTTGAAAGAGTTTGGAAACAGAGTTTCAACAGCAGAGTTGAA 356  
QY 307 GATGATGATAGCTCTGTTCTGAGAGAGTCTAGTAATGACTCGAAGATTAATCCACA 366  
DB 357 GATGATGATAGCTCTGTTCTGAGAGAGTCTAGTAATGACTCGAAGATTAATCCACA 416  
QY 367 CGGAACAGAGGTTTCCAGAGAGCGGCTATCGAGATGGAATTAATCA- GAAGCTTC 425  
DB 417 CGGAACAGAGGTTTCCAGAGAGCGGCTATCGAGATGGAATTAATCA- GAAGCTTC 476

QY 426 AGGCCATACAGAGAGTGAAGAGTAGTTCCAGAGTTCCCGTGAGAGATTGGTCT 485  
DB 477 AGGCCATACAGAGAGTGAAGAGTAGTTCCAGAGTTCCCGTGAGAGATTGGTCT 536  
QY 486 AGGAGTCCAAATATGACTTACAGCCAGCAAGTATGACAGCGACTGGCTTTT 545  
DB 537 AGGAGTCCAAATATGACTTACAGCCAGCAAGTATGACAGCGACTGGCTTTT 596  
QY 546 TGGTTTAAAGACCACTTATAGTGGACAGTATGTTGATCTTCCAAAGCAAG 605  
DB 597 TGGTTTAAAGACCACTTATAGTGGACAGTATGTTGATCTTCCAAAGCAAG 656  
QY 606 T 606  
DB 657 T 657  
RESULT 8  
EG717354  
LOCUS 651 bp mRNA linear EST 08-MAY-2001  
DEFINITION 60268968F1 NIH\_MGC\_97 Homo sapiens CDNA clone IMAGE:482149 5',  
mRNA sequence.  
ACCESSION EG717354  
VERSION EG717354  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 651)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM10729 row: D column: 22  
High quality sequence stop: 651.  
Location/Qualifiers  
1. 651  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1ib="IMAGE:482149"  
/clone\_1ib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gcgcag  
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTN-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 194 a 105 c 195 g 157 t  
ORIGIN  
Query Match 21.6%; Score 469; DB 12; Length 651;  
Best Local Similarity 100.0%; Pred. No. 3.4e-245;  
Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 16 TGGGAAGCAGAATCAACCTCATATGTTCTTATGTTCCCATATTGAGAGATAG 75  
DB 77 TGGGAAGCAGAATCAACCTCATATGTTCTTATGTTCCCATATTGAGAGATAG 136



```

/clone_1lb="NIH_MGC_97"
/lab_host="DH10B"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

```

```

BASE COUNT      251 a      153 c      248 g      194 t
ORIGIN

```

```

Query Match      20.3%; Score 440; DB 12; Length 846;
Best Local Similarity 100.0%; Pred. No. 3e-229;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 GATGAAGATTGGAGACGAAATCAACCTCATATGCTTCTCATATGTTCCCATATTTGAG 66
DB 78 GATGAAGATTGGAGACGAAATCAACCTCATATGCTTCTCATATGTTCCCATATTTGAG 137
QY 67 AAGGATAGGATTTCTGAGAAATGAGAGCAATTTTAAACAGACTCCAGCTTCATATCA 126
DB 138 AAGGATAGGATTTCTGAGAAATGAGAGCAATTTTAAACAGACTCCAGCTTCATATCA 197
QY 127 GAATGATGATGACCTTCTCGAAGAGATCATTTCTGAAAGTGATTTGCTCTGGG 186
DB 198 GAATGATGATGACCTTCTCGAAGAGATCATTTCTGAAAGTGATTTGCTCTGGG 257
QY 187 CGGAATTTTGAACAGAGATGCTGTGATGTAATAGCAGATATATACATCCACATG 246
DB 258 CGGAATTTTGAACAGAGATGCTGTGATGTAATAGCAGATATATACATCCACATG 317
QY 247 GGTGCTTTTGAAGTGGAAAGAGTTTGGAAACAGAGCTTTTCAACAGCAGCTTTGAA 306
DB 318 GGTGCTTTTGAAGTGGAAAGAGTTTGGAAACAGAGCTTTTCAACAGCAGCTTTGAA 377
QY 307 GATGATGATAGCTCTGCTTCTGAGAGAGTCTAGTAATAGCTCGAAGATTAATCCACA 366
DB 378 GATGATGATAGCTCTGCTTCTGAGAGAGTCTAGTAATAGCTCGAAGATTAATCCACA 437
QY 367 CGGAACAGAGGGTTTTCAGAGAGCGGCTATCGAGATGAAATATATTCAGAGCTTCA 426
DB 438 CGGAACAGAGGGTTTTCAGAGAGCGGCTATCGAGATGAAATATATTCAGAGCTTCA 497
QY 427 GGGCCATACAGAGAGGTGG 446
DB 498 GGGCCATACAGAGAGGTGG 517

```

```

RESULT 11
LOCUS      BI461393      861 bp      mRNA      linear      EST 21-AUG-2001
DEFINITION 603206861.F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272430 5',
ACCESSION  BI461393
VERSION    BI461393.1 GI:15252049
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 861)
           NIH-MGC http://mgs.nci.nih.gov/
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgaps-r@mail.nih.gov
           Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
           cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
           Toshiyuki and Piero Carninci (RIKEN)
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
COMMENT

```

```

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
plate: LLML1687 row: j column: 15
High quality sequence stop: 672.
Location/Qualifiers
1. 861

```

```

FEATURES
source
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:5272430"
  /clone_1lb="NIH_MGC_97"
  /lab_host="DH10B"

```

```

/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

```

```

BASE COUNT      265 a      147 c      247 g      201 t
ORIGIN

```

```

Query Match      19.8%; Score 430; DB 13; Length 861;
Best Local Similarity 99.7%; Pred. No. 9.3e-224;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 7 GATGAAGATTGGAGACGAAATCAACCTCATATGCTTCTCATATGTTCCCATATTTGAG 66
DB 55 GATGAAGATTGGAGACGAAATCAACCTCATATGCTTCTCATATGTTCCCATATTTGAG 114
QY 67 AAGGATAGGATTTCTGAGAAATGAGAGCAATTTTAAACAGACTCCAGCTTCATATCA 126
DB 115 AAGGATAGGATTTCTGAGAAATGAGAGCAATTTTAAACAGACTCCAGCTTCATATCA 174
QY 127 GAATGATGATGACCTTCTCGAAGAGATCATTTCTGAAAGTGATTTGCTCTGGG 186
DB 175 GAATGATGATGACCTTCTCGAAGAGATCATTTCTGAAAGTGATTTGCTCTGGG 234
QY 187 CGGAATTTTGAACAGAGATGCTGTGATGTAATAGCAGATATATACATCCACATG 246
DB 235 CGGAATTTTGAACAGAGATGCTGTGATGTAATAGCAGATATATACATCCACATG 294
QY 247 GGTGCTTTTGAAGTGGAAAGAGTTTGGAAACAGAGCTTTTCAACAGCAGCTTTGAA 306
DB 295 GGTGCTTTTGAAGTGGAAAGAGTTTGGAAACAGAGCTTTTCAACAGCAGCTTTGAA 354
QY 307 GATGATGATAGCTCTGCTTCTGAGAGAGTCTAGTAATAGCTCGAAGATTAATCCACA 366
DB 355 GATGATGATAGCTCTGCTTCTGAGAGAGTCTAGTAATAGCTCGAAGATTAATCCACA 414
QY 367 CGGAACAGAGGGTTTTCAGAGAGCGGCTATCGAGATGAAATATATTCAGAGCTTCA 426
DB 415 CGGAACAGAGGGTTTTCAGAGAGCGGCTATCGAGATGAAATATATTCAGAGCTTCA 474
QY 427 GGGCCATACAGAGAGGTGGAGAGTATTTCCAGAGCTTCCGCTGAGAGATTTGCTCTA 466
DB 475 GGGCCATACAGAGAGGTGGAGAGTATTTCCAGAGCTTCCGCTGAGAGATTTGCTCTA 533
QY 487 GGAAGTCCAATATAGCTTTCAGAGAGCGGCTATCGAGATGAAATATATTCAGAGCTTCA 546
DB 534 GGAAGTCCAATATATAGCTTTCAGAGAGCGGCTATCGAGATGAAATATATTCAGAGCTTCA 593
QY 547 GGTCTAGAGAGCAGATATTAAGTGCACAGGATATGATGATCTCTCAACAGAGAGT 606
DB 594 GGTCTAGAGAGCAGATATTAAGTGCACAGGATATGATGATCTCTCAACAGAGAGT 653
QY 607 GG 608
DB 654 GG 655

```

RESULT 12  
BI462207 830 bp mRNA linear EST 21-AUG-2001  
LOCUS 603205520F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5271221 5',  
DEFINITION mRNA sequence.  
ACCESSION BI462207 GI:15252863  
VERSION BI462207  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 830)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Miklos Palokvits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11684 row: h column: 06  
High quality sequence stop: 769.  
Location/Qualifiers  
1. 830  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5271221"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTCTTTTCTTTTNA-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 255 a 137 c 244 g 193 t 1 others  
ORIGIN  
Query Match 19.4%; Score 422; DB 13; Length 830;  
Best Local Similarity 99.8%; Pred. No. 2.3e-219;  
Matches 472; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 7 GATGAATTTGGGAAGCAATCAACCTCATATGCTTCTTCTTCCCATTTTGGAG 66  
DB 78 GATGAATTTGGGAAGCAATCAACCTCATATGCTTCTTCTTCCCATTTTGGAG 137  
QY 67 AAGGATAGGATTTCTGGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTCATCA 126  
DB 138 AAGGATAGGATTTCTGGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTCATCA 197  
QY 127 GAAATGATGATGACCTTCTCGAAGAGATCAATTTCAAGAAAGTGGATTGGCTGGG 186  
DB 198 GAAATGATGATGACCTTCTCGAAGAGATCAATTTCAAGAAAGTGGATTGGCTGGG 257  
QY 187 CGGAATTTTGGAAACAGAGATGCTGTGAGTGAATAAGGAGATTAATACATCCACATG 246  
DB 258 CGGAATTTTGGAAACAGAGATGCTGTGAGTGAATAAGGAGATTAATACATCCACATG 317  
QY 247 GGTGCTTTTGGAGTTGGAAAGAGTTTGGAAACAGAGAGTTTTCAGACAGAGTTTGA 306  
DB 318 GGTGCTTTTGGAGTTGGAAAGAGTTTGGAAACAGAGAGTTTTCAGACAGAGTTTGA 377  
QY 307 GATGATGATGATGATGCTTCTCGAAGAGATCTAGTAATGATGCGAAGATTAATCCA 366

DB 378 GATGATGATGATGCTTCTCGAAGAGATCTAGTAATGATGCGAAGATTAATCCA 437  
QY 367 CGGAACAGAGAGTTTTCAGAGAGAGGCGGTATCCAGATGAAATATTCAGAGCTTCA 426  
DB 438 CGGAACAGAGAGTTTTCAGAGAGAGGCGGTATCCAGATGAAATATTCAGAGCTTCA 497  
QY 427 GGGCCATACAGAGAGAGTGAAGAGTGAAGTTCGAGTTGCCGTGAGAGATT 479  
DB 498 GGGCCATACAGAGAGAGTGAAGAGTGAAGTTCGAGTTGCCGTGAGAGATT 550  
RESULT 13  
BG771822 699 bp mRNA linear EST 15-MAY-2001  
LOCUS 602720504F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4837296 5',  
DEFINITION mRNA sequence.  
ACCESSION BG771822 GI:14082475  
VERSION BG771822  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 699)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Miklos Palokvits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10769 row: h column: 01  
High quality sequence stop: 692.  
Location/Qualifiers  
1. 699  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4837296"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTCTTTTCTTTTNA-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 208 a 113 c 211 g 167 t  
ORIGIN  
Query Match 19.0%; Score 413; DB 12; Length 699;  
Best Local Similarity 99.8%; Pred. No. 1.8e-214;  
Matches 533; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 7 GATGAATTTGGGAAGCAATCAACCTCATATGCTTCTTCTTCCCATTTTGGAG 66  
DB 84 GATGAATTTGGGAAGCAATCAACCTCATATGCTTCTTCTTCCCATTTTGGAG 143  
QY 67 AAGGATAGGATTTCTGGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTCATCA 126  
DB 144 AAGGATAGGATTTCTGGAGAAATGAGACAAATTTTAAAGAGCTCCAGCTTCATCA 203  
QY 127 GAAATGATGATGACCTTCTCGAAGAGATCAATTTCAAGAAAGTGGATTGGCTGGG 186

Db 204 GAAATGATGATGACCTTCTCGAAGATCATTTATGAAAAGTGATTCCTCTGGG 263  
Qy 187 CGGAATTTTGAACAAGATGCTGTGATTAATAGCAGATATATACATCAATG 246  
Db 264 CGGAATTTTGAACAAGATGCTGTGATTAATAGCAGATATATACATCAATG 323  
Qy 247 GGTGGTTTGGAGTTGGAAGAAGTTTGGAAACAGAGTTTTCACACAGCTTTGAA 306  
Db 324 GTTGGG-TTTGGAGTTGGAAGAAGTTTGGAAACAGAGTTTTCACACAGCTTTGAA 382  
Qy 307 GATGGATAGCTCTGCTGTTCTGAGAGAGTCAATATAGCTCGAAGATATCCACA 366  
Db 383 GATGGATAGCTCTGCTGTTCTGAGAGAGTCAATATAGCTCGAAGATATCCACA 442  
Qy 367 CGGAAGAGAGGTTTTCAGAGAGAGGCGCTATCGAGATGAATATATTCAGAGCTTCA 426  
Db 443 CGGAAGAGAGGTTTTCAGAGAGAGGCGCTATCGAGATGAATATATTCAGAGCTTCA 502  
Qy 427 GGGCCATACAGAAAGAGTGAAGAGTATGTTCCGAGTTCCCGTGAAGATTTGGCTTA 486  
Db 503 GGGCCATACAGAAAGAGTGAAGAGTATGTTCCGAGTTCCCGTGAAGATTTGGCTTA 562  
Qy 487 GGAAGTCCAAATATATGATTAAGCCAGAGATATATGAGGCACTGGTGGC 540  
Db 563 GGAAGTCCAAATATATGATTAAGCCAGAGATATATGAGGCACTGGTGGC 616

RESULT 14  
LOCUS B1463502 743 bp mRNA linear EST 21-AUG-2001  
DEFINITION 603203596F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5269511 5',  
mRNA sequence.  
ACCESSION B1463502  
VERSION B1463502.1 GI:15254145  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM11679 row: p column: 24  
High quality sequence stop: 738.  
Location/Qualifiers

## FEATURES

source

1. 743  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5269511"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-xhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVA-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 227 a 116 c 222 g 177 t 1 others  
ORIGIN

Query Match 19.0%; Score 412; DB 13; Length 743;  
Best Local Similarity 99.8%; Pred. No. 6,6e-214;  
Matches 462; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 GGGAGCAGAAATCAACCTCATATGTCTTCTATGTCCCATATTTGAGAGATAGT 76  
Db 98 GGGAGCAGAAATCAACCTCATATGTCTTCTATGTCCCATATTTGAGAGATAGT 157  
Qy 77 ATCTGAGAGAAATGAGCAATTTTAACAGAGCTCAGCTTCATCATCAGAAATGAGT 136  
Db 158 ATCTGAGAGAAATGAGCAATTTTAACAGAGCTCAGCTTCATCATCAGAAATGAGT 217  
Qy 137 ATGACCTTCTCGAAGATCATTTTCATGAAAAGTGATTTGCTGCGGCGAATTTTG 196  
Db 218 ATGACCTTCTCGAAGATCATTTTCATGAAAAGTGATTTGCTGCGGCGAATTTTG 277  
Qy 197 GAAACAGATGCTGTGATGTAATTAAGCAGATATATCATCCAAATGGTGTGTTG 256  
Db 278 GAAACAGATGCTGTGATGTAATTAAGCAGATATATCATCCAAATGGTGTGTTG 337  
Qy 257 GAGTTGAAAAGGTTTGGAAACAGAGTTTTCAAACAGAGGTTTGAAGATGGATA 316  
Db 338 GAGTTGAAAAGGTTTGGAAACAGAGTTTTCAAACAGAGGTTTGAAGATGGATA 397  
Qy 317 GCTCTGTTCTGAGAGAGTCTAGTAATGACTCGAAGATATTCACACGAGACAG 376  
Db 398 GCTCTGTTCTGAGAGAGTCTAGTAATGACTCGAAGATATTCACACGAGACAG 457  
Qy 377 GGTTCCTCAAGAGAGCGGCTATCGAGATGAATATTCAGAGCTTCAGGCGCATACA 436  
Db 458 GGTTCCTCAAGAGAGCGGCTATCGAGATGAATATTCAGAGCTTCAGGCGCATACA 517  
Qy 437 GAAGAGTGAAGAGTATGTTCCGAGTTCCCGTGAAGATT 479  
Db 518 GAAGAGTGAAGAGTATGTTCCGAGTTCCCGTGAAGATT 560

RESULT 15  
LOCUS B1459115 841 bp mRNA linear EST 21-AUG-2001  
DEFINITION 603199485F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5265501 5',  
mRNA sequence.  
ACCESSION B1459115  
VERSION B1459115.1 GI:15249771  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM11669 row: i column: 22  
High quality sequence stop: 646.  
Location/Qualifiers

## FEATURES

source

1. 841  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5265501"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"

/note="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site 1: BamHI; Site 2: SalI-xhoI (Gtcgag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.2 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIMH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 235 a 175 c 244 g 187 t  
ORIGIN

Query Match 18.0%; Score 392; DB 13; Length 841;  
Best Local Similarity 99.8%; Pred. No. 6.5e-203;  
Matches 512; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 16 TGGGAACGACAAATCAACCTCATATGTCTTCTCATATGTTCCCATATTTGAGAGATAGG 75  
DB 96 TGGGAACGACAAATCAACCTCATATGTCTTCTCATATGTTCCCATATTTGAGAGATAGG 155  
QY 76 TATCTGAGAAAAATGAGACAATTTTAACAGACCTCAGCTTCATCATCAGAAATGAT 135  
DB 156 TATCTGAGAAAAATGAGACAATTTTAACAGACCTCAGCTTCATCATCAGAAATGAT 215  
QY 136 GATGACCTTCTCGAAGAGATCATTTTCATGAAAAAGTATTTGCTCTGGCGGAAATTT 195  
DB 216 GATGACCTTCTCGAAGAGATCATTTTCATGAAAAAGTATTTGCTCTGGCGGAAATTT 275  
QY 196 GGAACGAGATGCTGTGATGTAATAGCGAGATATATCATCCAAATGGGTGTTT 255  
DB 276 GGAACGAGATGCTGTGATGTAATAGCGAGATATATCATCCAAATGGGTGTTT 335  
QY 256 GGAATTTGAA-AGAGTTTGGAAACAGAGTTTTCAAACAGAGTTTGAAGTGTGA 314  
DB 336 GGAATTTGAAACAGAGTTTTCAAACAGAGTTTGAAGTGTGA 395  
QY 315 TAGCTCTGTTCTGAGAGAGTCTAGTAATGACTGCAAGATATATCAACACGAAACAG 374  
DB 396 TAGCTCTGTTCTGAGAGAGTCTAGTAATGACTGCAAGATATATCAACACGAAACAG 455  
QY 375 AGGTTTTCAGAGAGCGGCTATCGAGATGAAATATTTCAAGACTTCAGGCGCAT 434  
DB 456 AGGTTTTCAGAGAGCGGCTATCGAGATGAAATATTTCAAGACTTCAGGCGCAT 515  
QY 435 CAGAGAGGTGAGAGAGTATCCGAGTTGCCGTGAGAGATTTGGTCTAGGAAAGTCC 494  
DB 516 CAGAGAGGTGAGAGAGTATCCGAGTTGCCGTGAGAGATTTGGTCTAGGAAAGTCC 575  
QY 495 AATATGACTTAGACCCAGACGATGTATGCA 527  
DB 576 AATATGACTTAGACCCAGACGATGTATGCA 608

Search completed: June 10, 2003, 14:51:37  
Job time: 2833 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 14:05:25 ; Search time 296 Seconds  
(without alignments)  
10269.578 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 2172  
Sequence: 1 atggggagatgaagatcggga.....tagatgatgagtcgcatggaat 2172Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 870385 seqs, 699768693 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PC7\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	340	15.7	420	10 US-09-833-381-1475	Sequence 1475, Ap
2	255	11.7	591	9 US-09-764-891-1170	Sequence 1170, Ap
3	220	10.1	2194	9 US-09-764-891-7481	Sequence 7481, Ap
4	220	10.1	2195	9 US-09-764-891-7482	Sequence 7482, Ap
5	220	10.1	2197	9 US-09-764-891-7483	Sequence 7483, Ap
6	21	1.0	1842	10 US-09-815-242-6893	Sequence 6893, Ap
7	21	1.0	1830121	9 US-10-329-960-1	Sequence 1, Appl
8	20	0.9	244	10 US-09-878-574-10286	Sequence 10286, A
9	20	0.9	266	10 US-09-878-574-12453	Sequence 12453, A
10	19	0.9	138	10 US-09-864-761-28835	Sequence 28835, A
11	19	0.9	418	9 US-09-918-995-35717	Sequence 35717, A
12	19	0.9	592	10 US-09-925-299-60	Sequence 12255, A
13	19	0.9	1472	10 US-09-925-299-60	Sequence 60, Appl
14	19	0.9	1472	10 US-09-925-301-185	Sequence 185, Appl
15	19	0.9	1987	10 US-09-925-301-185	Sequence 480, Appl
16	19	0.9	2451	12 US-10-044-090-480	Sequence 84, Appl
17	19	0.9	4250	9 US-09-957-187-84	Sequence 67, Appl
18	19	0.9	4423	12 US-10-002-600-67	Sequence 1, Appl
19	19	0.9	536165	9 US-09-939-964-1	

20	18	0.8	297	10 US-09-923-876-3712	Sequence 3712, Ap
21	18	0.8	379	9 US-09-991-936-1246	Sequence 1246, Ap
22	18	0.8	381	9 US-09-991-936-725	Sequence 725, Ap
23	18	0.8	413	9 US-09-918-995-7261	Sequence 7261, Ap
24	18	0.8	421	9 US-10-076-622-116	Sequence 116, Ap
25	18	0.8	421	10 US-09-604-287A-116	Sequence 116, Ap
26	18	0.8	421	10 US-09-339-338-116	Sequence 116, Ap
27	18	0.8	421	12 US-10-007-805-116	Sequence 116, Ap
28	18	0.8	423	10 US-09-864-761-11121	Sequence 11121, A
29	18	0.8	455	9 US-09-918-995-11308	Sequence 11308, A
30	18	0.8	478	9 US-09-918-995-20686	Sequence 20686, A
31	18	0.8	511	10 US-09-783-590-183	Sequence 183, Ap
32	18	0.8	532	10 US-09-755-325-3	Sequence 3, Appl
33	18	0.8	587	9 US-09-332-785-336	Sequence 336, Ap
34	18	0.8	598	10 US-09-833-381-99	Sequence 99, Appl
35	18	0.8	599	9 US-09-854-531-816	Sequence 816, Ap
36	18	0.8	606	9 US-10-198-846-7904	Sequence 7904, Ap
37	18	0.8	1205	9 US-09-533-029-5	Sequence 5, Appl
38	18	0.8	1536	10 US-09-755-325-5	Sequence 2560, Ap
39	18	0.8	1614	9 US-09-938-842A-2560	Sequence 7, Appl
40	18	0.8	1682	10 US-09-755-325-7	Sequence 19, Appl
41	18	0.8	1914	9 US-10-051-902-19	Sequence 19, Appl
42	18	0.8	1914	9 US-10-051-909-19	Sequence 16, Appl
43	18	0.8	2005	12 US-09-938-842A-3617	Sequence 1, Appl
44	18	0.8	2005	9 US-10-142-373-1	Sequence 15, Appl
45	18	0.8	2196	9 US-10-086-464-16	Sequence 15, Appl
46	18	0.8	2261	9 US-10-086-464-15	Sequence 5, Appl
47	18	0.8	2275	9 US-09-825-486-3	Sequence 3, Appl
48	18	0.8	3103	8 US-08-870-434-2	Sequence 2, Appl
49	18	0.8	3103	9 US-09-560-150-2	Sequence 2, Appl
50	18	0.8	3103	9 US-10-067-741-2	Sequence 2, Appl
51	18	0.8	3103	10 US-09-372-044-3	Sequence 3, Appl
52	18	0.8	3103	10 US-09-924-417-66	Sequence 66, Appl
53	18	0.8	3111	9 US-10-181-603-3	Sequence 3, Appl
54	18	0.8	3111	10 US-09-954-456-1877	Sequence 825, Ap
55	18	0.8	9551	9 US-09-764-872-825	Sequence 1330, Ap
56	18	0.8	13467	9 US-09-764-868-1330	Sequence 1498, Ap
57	18	0.8	13467	9 US-09-764-868-1498	Sequence 8632, Ap
58	18	0.8	13548	9 US-09-764-891-8632	Sequence 823, Ap
59	18	0.8	13548	9 US-09-764-872-823	Sequence 1018, Ap
60	18	0.8	18385	9 US-10-074-095-1018	Sequence 325, Ap
61	18	0.8	18385	10 US-09-764-860-1018	Sequence 823, Ap
62	18	0.8	20020	9 US-10-091-483-325	Sequence 902, Ap
63	18	0.8	20020	9 US-09-764-891-8233	Sequence 4312, Ap
64	18	0.8	20020	10 US-09-764-846-325	Sequence 189, Ap
65	18	0.8	24	9 US-09-940-185-902	Sequence 31834, A
66	17	0.8	37	9 US-09-747-377-189	Sequence 27937, A
67	17	0.8	54	9 US-09-747-377-189	Sequence 9249, Ap
68	17	0.8	54	9 US-09-747-377-189	Sequence 8481, Ap
69	17	0.8	135	10 US-09-864-761-31834	Sequence 4334, Ap
70	17	0.8	135	10 US-09-864-761-27937	Sequence 408, Ap
71	17	0.8	155	10 US-09-864-761-19712	Sequence 1436, Ap
72	17	0.8	168	10 US-09-878-574-9249	Sequence 199, Ap
73	17	0.8	178	10 US-09-878-574-8481	Sequence 361, Ap
74	17	0.8	201	10 US-09-728-446-581	Sequence 3707, A
75	17	0.8	278	10 US-09-294-093B-4334	Sequence 9239, Ap
76	17	0.8	308	9 US-10-040-739-408	Sequence 1610, Ap
77	17	0.8	318	9 US-09-738-626-1436	Sequence 6299, Ap
78	17	0.8	318	9 US-09-764-872-199	Sequence 2931, Ap
79	17	0.8	320	9 US-09-864-761-2931	Sequence 8795, Ap
80	17	0.8	370	10 US-09-918-995-3707	Sequence 9257, Ap
81	17	0.8	384	9 US-09-960-335-9239	Sequence 10666, A
82	17	0.8	392	10 US-09-960-335-9239	Sequence 2245, Ap
83	17	0.8	395	9 US-10-066-543-1610	Sequence 310, Ap
84	17	0.8	421	10 US-09-867-701-6299	
85	17	0.8	421	10 US-09-864-761-2931	
86	17	0.8	424	10 US-09-878-574-5113	
87	17	0.8	434	10 US-09-960-335-8795	
88	17	0.8	450	10 US-09-960-335-9252	
89	17	0.8	464	10 US-09-867-701-10666	
90	17	0.8	464	10 US-09-864-761-2245	
91	17	0.8	472	10 US-09-864-761-310	
92	17	0.8	472	10 US-09-864-761-14159	

93	17	0.8	476	10	US-09-864-761-11357	Sequence 11357, A	C 166	17	0.8	3170	9	US-10-174-572-249	Sequence 249, App
94	17	0.8	479	9	US-09-918-995-34562	Sequence 34562, A	C 167	17	0.8	3170	9	US-10-174-579-249	Sequence 249, App
C 95	17	0.8	489	9	US-10-046-935-7	Sequence 7, Appl	C 168	17	0.8	3170	9	US-10-174-582-249	Sequence 249, App
C 96	17	0.8	489	9	US-09-878-178-7	Sequence 7, Appl	C 169	17	0.8	3170	9	US-10-174-588-249	Sequence 249, App
C 97	17	0.8	489	9	US-10-146-502-7	Sequence 2007, Ap	C 170	17	0.8	3170	9	US-10-175-739-249	Sequence 249, App
C 98	17	0.8	490	10	US-09-867-550-2007	Sequence 2007, Ap	C 171	17	0.8	3170	9	US-10-175-740-249	Sequence 249, App
C 99	17	0.8	500	10	US-09-827-998-52	Sequence 528, Appl	C 172	17	0.8	3170	9	US-10-175-743-249	Sequence 249, App
C 100	17	0.8	502	10	US-09-783-590-238	Sequence 238, Appl	C 173	17	0.8	3170	9	US-10-176-488-249	Sequence 249, App
C 101	17	0.8	512	10	US-09-864-761-15312	Sequence 15312, A	C 174	17	0.8	3170	9	US-10-176-492-249	Sequence 249, App
C 102	17	0.8	518	9	US-10-056-543-1729	Sequence 1729, Ap	C 175	17	0.8	3170	9	US-10-176-747-249	Sequence 249, App
C 103	17	0.8	537	9	US-09-972-473-12	Sequence 12, Appl	C 176	17	0.8	3170	9	US-10-176-750-249	Sequence 249, App
C 104	17	0.8	571	9	US-10-046-935-171	Sequence 171, App	C 177	17	0.8	3170	9	US-10-176-985-249	Sequence 249, App
C 105	17	0.8	571	9	US-09-878-178-171	Sequence 171, App	C 178	17	0.8	3170	9	US-10-176-987-249	Sequence 249, App
C 106	17	0.8	571	9	US-10-146-502-171	Sequence 171, App	C 179	17	0.8	3170	9	US-10-176-991-249	Sequence 249, App
C 107	17	0.8	593	9	US-10-066-543-235	Sequence 235, App	C 180	17	0.8	3170	9	US-10-176-992-249	Sequence 249, App
C 108	17	0.8	624	9	US-09-976-736-6	Sequence 6, Appl	C 181	17	0.8	3170	9	US-10-176-993-249	Sequence 249, App
C 109	17	0.8	697	9	US-09-774-639-46	Sequence 46, Appl	C 182	17	0.8	3170	9	US-10-184-658-249	Sequence 249, App
C 110	17	0.8	697	9	US-09-969-730-32	Sequence 32, Appl	C 183	17	0.8	3170	9	US-10-173-695-249	Sequence 249, App
C 111	17	0.8	702	9	US-09-972-473-10	Sequence 10, Appl	C 184	17	0.8	3170	9	US-10-173-697-249	Sequence 249, App
C 112	17	0.8	720	9	US-09-764-891-6426	Sequence 6426, Ap	C 185	17	0.8	3170	9	US-10-173-705-249	Sequence 249, App
C 113	17	0.8	744	10	US-09-815-242-9363	Sequence 9363, Ap	C 186	17	0.8	3170	9	US-10-174-576-249	Sequence 249, App
C 114	17	0.8	744	10	US-09-815-242-9524	Sequence 9524, Ap	C 187	17	0.8	3170	9	US-10-174-585-249	Sequence 249, App
C 115	17	0.8	780	9	US-09-976-736-5	Sequence 5, Appl	C 188	17	0.8	3170	9	US-10-174-586-249	Sequence 249, App
C 116	17	0.8	789	9	US-09-972-473-22	Sequence 22, Appl	C 189	17	0.8	3170	9	US-10-175-747-249	Sequence 249, App
C 117	17	0.8	828	9	US-09-764-872-851	Sequence 851, Appl	C 190	17	0.8	3170	9	US-10-176-481-249	Sequence 249, App
C 118	17	0.8	828	9	US-09-764-872-852	Sequence 852, App	C 191	17	0.8	3170	9	US-10-176-485-249	Sequence 249, App
C 119	17	0.8	828	9	US-09-764-872-853	Sequence 853, App	C 192	17	0.8	3170	9	US-10-176-487-249	Sequence 249, App
C 120	17	0.8	828	9	US-09-764-872-854	Sequence 854, App	C 193	17	0.8	3170	9	US-10-176-493-249	Sequence 249, App
C 121	17	0.8	884	9	US-10-002-344-87	Sequence 87, Appl	C 194	17	0.8	3170	9	US-10-176-756-249	Sequence 249, App
C 122	17	0.8	915	9	US-10-115-928-26	Sequence 26, Appl	C 195	17	0.8	3170	9	US-10-176-911-249	Sequence 249, App
C 123	17	0.8	931	9	US-09-938-842A-2975	Sequence 2975, Ap	C 196	17	0.8	3170	9	US-10-176-919-249	Sequence 249, App
C 124	17	0.8	948	9	US-09-968-436B-5	Sequence 5, Appl	C 197	17	0.8	3170	9	US-10-176-925-249	Sequence 249, App
C 125	17	0.8	987	10	US-09-974-300-1852	Sequence 1852, Ap	C 198	17	0.8	3170	9	US-10-176-978-249	Sequence 249, App
C 126	17	0.8	1182	9	US-09-738-626-773	Sequence 773, App	C 199	17	0.8	3170	9	US-10-179-510-249	Sequence 249, App
C 127	17	0.8	1206	9	US-09-738-626-1966	Sequence 1666, Ap	C 200	17	0.8	3170	9	US-10-180-543-249	Sequence 249, App
C 128	17	0.8	1239	9	US-09-938-842A-1874	Sequence 1874, Ap	C 201	17	0.8	3170	9	US-10-180-544-249	Sequence 249, App
C 129	17	0.8	1401	9	US-09-935-371-32	Sequence 32, Appl	C 202	17	0.8	3170	9	US-10-180-546-249	Sequence 249, App
C 130	17	0.8	1446	10	US-09-884-260A-8	Sequence 8, Appl	C 203	17	0.8	3170	9	US-10-180-547-249	Sequence 249, App
C 131	17	0.8	1485	9	US-09-935-371-33	Sequence 33, Appl	C 204	17	0.8	3170	9	US-10-180-549-249	Sequence 249, App
C 132	17	0.8	1505	10	US-09-796-766-9	Sequence 9, Appl	C 205	17	0.8	3170	9	US-10-180-555-249	Sequence 249, App
C 133	17	0.8	1517	10	US-09-796-766-7	Sequence 7, Appl	C 206	17	0.8	3170	9	US-10-180-559-249	Sequence 249, App
C 134	17	0.8	1521	10	US-09-801-568-89	Sequence 89, Appl	C 207	17	0.8	3170	9	US-10-181-000-249	Sequence 249, App
C 135	17	0.8	1542	10	US-09-887-576-230	Sequence 230, App	C 208	17	0.8	3170	9	US-10-183-010-249	Sequence 249, App
C 136	17	0.8	1674	9	US-09-764-891-8862	Sequence 8862, Ap	C 209	17	0.8	3170	9	US-10-183-012-249	Sequence 249, App
C 137	17	0.8	1674	9	US-09-764-891-8865	Sequence 8865, Ap	C 210	17	0.8	3170	9	US-10-184-614-249	Sequence 249, App
C 138	17	0.8	1715	10	US-09-884-260A-56	Sequence 56, Appl	C 211	17	0.8	3170	9	US-10-184-623-249	Sequence 249, App
C 139	17	0.8	1869	9	US-09-935-371-40	Sequence 40, Appl	C 212	17	0.8	3170	9	US-10-184-635-249	Sequence 249, App
C 140	17	0.8	1906	9	US-10-218-547-15	Sequence 15, Appl	C 213	17	0.8	3170	9	US-10-184-637-249	Sequence 249, App
C 141	17	0.8	1924	10	US-09-765-272-159	Sequence 159, App	C 214	17	0.8	3170	9	US-10-184-646-249	Sequence 249, App
C 142	17	0.8	2000	9	US-09-938-842A-4042	Sequence 4042, Ap	C 215	17	0.8	3170	9	US-10-184-647-249	Sequence 249, App
C 143	17	0.8	2000	9	US-09-938-842A-4108	Sequence 4108, Ap	C 216	17	0.8	3170	9	US-10-184-652-249	Sequence 249, App
C 144	17	0.8	2000	9	US-09-938-842A-4329	Sequence 4329, Ap	C 217	17	0.8	3170	9	US-10-187-594-249	Sequence 249, App
C 145	17	0.8	2000	9	US-09-938-842A-4329	Sequence 4329, Ap	C 218	17	0.8	3170	9	US-10-187-596-249	Sequence 249, App
C 146	17	0.8	2000	9	US-09-938-842A-5152	Sequence 5152, Ap	C 219	17	0.8	3170	9	US-10-187-745-249	Sequence 249, App
C 147	17	0.8	2120	9	US-09-344-882-29	Sequence 29, Appl	C 220	17	0.8	3170	9	US-10-187-885-249	Sequence 249, App
C 148	17	0.8	2219	9	US-09-961-721-1	Sequence 1, Appl	C 221	17	0.8	3170	9	US-10-187-886-249	Sequence 249, App
C 149	17	0.8	2235	9	US-09-769-787-196	Sequence 196, App	C 222	17	0.8	3170	9	US-10-187-888-249	Sequence 249, App
C 150	17	0.8	2509	9	US-10-108-605-86	Sequence 86, Appl	C 223	17	0.8	3170	9	US-10-187-889-249	Sequence 249, App
C 151	17	0.8	2548	9	US-09-935-371-41	Sequence 41, Appl	C 224	17	0.8	3170	9	US-10-176-760-249	Sequence 249, App
C 152	17	0.8	2693	9	US-09-850-948-1	Sequence 1, Appl	C 225	17	0.8	3170	9	US-10-176-990-249	Sequence 249, App
C 153	17	0.8	3124	9	US-10-219-220-131	Sequence 131, Appl	C 226	17	0.8	3170	9	US-10-180-541-249	Sequence 249, App
C 154	17	0.8	3170	9	US-10-174-590-249	Sequence 249, App	C 227	17	0.8	3170	9	US-10-180-542-249	Sequence 249, App
C 155	17	0.8	3170	9	US-10-176-758-249	Sequence 249, App	C 228	17	0.8	3170	9	US-10-180-548-249	Sequence 249, App
C 156	17	0.8	3170	9	US-10-175-737-249	Sequence 249, App	C 229	17	0.8	3170	9	US-10-180-551-249	Sequence 249, App
C 157	17	0.8	3170	9	US-10-175-737-249	Sequence 249, App	C 230	17	0.8	3170	9	US-10-180-998-249	Sequence 249, App
C 158	17	0.8	3170	9	US-10-173-706-249	Sequence 249, App	C 231	17	0.8	3170	9	US-10-180-999-249	Sequence 249, App
C 159	17	0.8	3170	9	US-10-175-752-249	Sequence 249, App	C 232	17	0.8	3170	9	US-10-183-013-249	Sequence 249, App
C 160	17	0.8	3170	9	US-10-176-482-249	Sequence 249, App	C 233	17	0.8	3170	9	US-10-184-612-249	Sequence 249, App
C 161	17	0.8	3170	9	US-10-176-757-249	Sequence 249, App	C 234	17	0.8	3170	9	US-10-184-616-249	Sequence 249, App
C 162	17	0.8	3170	9	US-10-176-913-249	Sequence 249, App	C 235	17	0.8	3170	9	US-10-184-617-249	Sequence 249, App
C 163	17	0.8	3170	9	US-10-180-952-249	Sequence 249, App	C 236	17	0.8	3170	9	US-10-184-622-249	Sequence 249, App
C 164	17	0.8	3170	9	US-10-180-557-249	Sequence 249, App	C 237	17	0.8	3170	9	US-10-184-628-249	Sequence 249, App
C 165	17	0.8	3170	9	US-10-173-700-249	Sequence 249, App	C 238	17	0.8	3170	9	US-10-184-629-249	Sequence 249, App

C 239	17	0.8	3170	9	US-10-184-630-249	Sequence 249, App	C 312	17	0.8	3170	9	US-10-180-560-249	Sequence 249, App
C 240	17	0.8	3170	9	US-10-184-631-249	Sequence 249, App	C 313	17	0.8	3170	9	US-10-180-561-249	Sequence 249, App
C 241	17	0.8	3170	9	US-10-184-632-249	Sequence 249, App	C 314	17	0.8	3170	9	US-10-184-611-249	Sequence 249, App
C 242	17	0.8	3170	9	US-10-184-636-249	Sequence 249, App	C 315	17	0.8	3170	9	US-10-184-620-249	Sequence 249, App
C 243	17	0.8	3170	9	US-10-184-640-249	Sequence 249, App	C 316	17	0.8	3170	9	US-10-184-642-249	Sequence 249, App
C 244	17	0.8	3170	9	US-10-184-650-249	Sequence 249, App	C 317	17	0.8	3170	9	US-10-184-655-249	Sequence 249, App
C 245	17	0.8	3170	9	US-10-184-651-249	Sequence 249, App	C 318	17	0.8	3170	9	US-10-192-010-249	Sequence 249, App
C 246	17	0.8	3170	9	US-10-187-888-249	Sequence 249, App	C 319	17	0.8	3170	9	US-10-195-894-249	Sequence 249, App
C 247	17	0.8	3170	9	US-10-187-597-249	Sequence 249, App	C 320	17	0.8	3170	9	US-10-205-908-249	Sequence 249, App
C 248	17	0.8	3170	9	US-10-187-598-249	Sequence 249, App	C 321	17	0.8	3170	9	US-10-186-885-249	Sequence 249, App
C 249	17	0.8	3170	9	US-10-187-600-249	Sequence 249, App	C 322	17	0.8	3170	9	US-10-187-599-249	Sequence 249, App
C 250	17	0.8	3170	9	US-10-187-601-249	Sequence 249, App	C 323	17	0.8	3170	9	US-10-187-750-249	Sequence 249, App
C 251	17	0.8	3170	9	US-10-187-602-249	Sequence 249, App	C 324	17	0.8	3170	9	US-10-188-780-249	Sequence 249, App
C 252	17	0.8	3170	9	US-10-187-603-249	Sequence 249, App	C 325	17	0.8	3170	9	US-10-192-015-249	Sequence 249, App
C 253	17	0.8	3170	9	US-10-187-741-249	Sequence 249, App	C 326	17	0.8	3170	9	US-10-194-394-249	Sequence 249, App
C 254	17	0.8	3170	9	US-10-187-743-249	Sequence 249, App	C 327	17	0.8	3170	9	US-10-194-425-249	Sequence 249, App
C 255	17	0.8	3170	9	US-10-187-746-249	Sequence 249, App	C 328	17	0.8	3170	9	US-10-194-485-249	Sequence 249, App
C 256	17	0.8	3170	9	US-10-187-747-249	Sequence 249, App	C 329	17	0.8	3170	9	US-10-195-885-249	Sequence 249, App
C 257	17	0.8	3170	9	US-10-187-751-249	Sequence 249, App	C 330	17	0.8	3170	9	US-10-195-889-249	Sequence 249, App
C 258	17	0.8	3170	9	US-10-187-753-249	Sequence 249, App	C 331	17	0.8	3170	9	US-10-195-899-249	Sequence 249, App
C 259	17	0.8	3170	9	US-10-187-754-249	Sequence 249, App	C 332	17	0.8	3170	9	US-10-196-748-249	Sequence 249, App
C 260	17	0.8	3170	9	US-10-187-757-249	Sequence 249, App	C 333	17	0.8	3170	9	US-10-196-750-249	Sequence 249, App
C 261	17	0.8	3170	9	US-10-187-884-249	Sequence 249, App	C 334	17	0.8	3170	9	US-10-197-699-249	Sequence 249, App
C 262	17	0.8	3170	9	US-10-188-767-249	Sequence 249, App	C 335	17	0.8	3170	9	US-10-197-700-249	Sequence 249, App
C 263	17	0.8	3170	9	US-10-188-769-249	Sequence 249, App	C 336	17	0.8	3170	9	US-10-197-705-249	Sequence 249, App
C 264	17	0.8	3170	9	US-10-188-770-249	Sequence 249, App	C 337	17	0.8	3170	9	US-10-197-708-249	Sequence 249, App
C 265	17	0.8	3170	9	US-10-188-773-249	Sequence 249, App	C 338	17	0.8	3170	9	US-10-198-764-249	Sequence 249, App
C 266	17	0.8	3170	9	US-10-188-781-249	Sequence 249, App	C 339	17	0.8				



C 531	17	0.8	3170	9	US-10-201-327-249	Sequence 249, App	604	16	0.7	175	10	US-09-878-574-12932	Sequence 12932, A
C 532	17	0.8	3170	9	US-10-006-116A-69	Sequence 69, Appl	605	16	0.7	177	9	US-10-083-357-267	Sequence 267, App
C 533	17	0.8	3170	9	US-10-006-117A-69	Sequence 69, Appl	606	16	0.7	180	10	US-09-867-701-6800	Sequence 6800, Ap
C 534	17	0.8	3170	9	US-10-013-913A-69	Sequence 69, Appl	607	16	0.7	192	10	US-09-783-590-162	Sequence 162, App
C 535	17	0.8	3170	9	US-10-017-527A-69	Sequence 69, Appl	608	16	0.7	207	10	US-09-864-761-2411	Sequence 29411, A
C 536	17	0.8	3170	9	US-10-173-696-249	Sequence 249, App	609	16	0.7	224	10	US-09-876-889-105	Sequence 105, App
C 537	17	0.8	3170	9	US-10-183-003-249	Sequence 249, App	610	16	0.7	229	9	US-09-878-574-5644	Sequence 5644, Ap
C 538	17	0.8	3170	9	US-10-183-003-249	Sequence 249, App	611	16	0.7	234	9	US-10-025-380-209	Sequence 209, App
C 539	17	0.8	3170	9	US-10-125-923A-249	Sequence 249, App	612	16	0.7	234	10	US-09-923-217-209	Sequence 209, App
C 540	17	0.8	3170	9	US-10-176-491-249	Sequence 249, App	613	16	0.7	234	10	US-09-923-263-209	Sequence 209, App
C 541	17	0.8	3170	9	US-10-176-979-249	Sequence 249, App	614	16	0.7	235	10	US-09-923-876-1335	Sequence 1435, Ap
C 542	17	0.8	3170	9	US-10-187-592-249	Sequence 249, App	615	16	0.7	236	10	US-09-924-035A-153	Sequence 153, App
C 543	17	0.8	3170	9	US-10-007-194A-69	Sequence 69, Appl	616	16	0.7	237	9	US-09-938-842A-1450	Sequence 1450, Ap
C 544	17	0.8	3170	9	US-10-013-430A-69	Sequence 69, Appl	617	16	0.7	240	9	US-10-083-357-265	Sequence 265, App
C 545	17	0.8	3170	9	US-10-197-691-249	Sequence 249, App	618	16	0.7	247	10	US-09-864-761-30341	Sequence 30341, A
C 546	17	0.8	3170	9	US-10-198-771-249	Sequence 249, App	619	16	0.7	250	9	US-09-736-457-27	Sequence 27, Appl
C 547	17	0.8	3170	9	US-10-011-671A-69	Sequence 69, Appl	620	16	0.7	250	9	US-09-902-941-27	Sequence 27, Appl
C 548	17	0.8	3170	9	US-10-012-755A-69	Sequence 69, Appl	621	16	0.7	250	9	US-09-849-626-27	Sequence 27, Appl
C 549	17	0.8	3170	9	US-10-174-575A-249	Sequence 249, App	622	16	0.7	250	9	US-10-017-754-27	Sequence 27, Appl
C 550	17	0.8	3170	9	US-10-179-520-249	Sequence 249, App	623	16	0.7	254	10	US-09-880-107-1346	Sequence 1346, Ap
C 551	17	0.8	3170	9	US-10-201-325-249	Sequence 249, App	624	16	0.7	260	10	US-09-923-876-6003	Sequence 6003, App
C 552	17	0.8	3170	9	US-10-202-941-249	Sequence 249, App	625	16	0.7	265	10	US-09-923-876-738	Sequence 738, App
C 553	17	0.8	3170	9	US-10-205-910-249	Sequence 249, App	626	16	0.7	269	9	US-09-809-391-142	Sequence 142, App
C 554	17	0.8	3170	9	US-10-015-386A-69	Sequence 69, Appl	627	16	0.7	270	10	US-09-833-790-83	Sequence 83, Appl
C 555	17	0.8	3170	9	US-10-179-526-249	Sequence 249, App	628	16	0.7	270	10	US-09-833-790-377	Sequence 377, App
C 556	17	0.8	3422	9	US-09-764-891-6046	Sequence 6046, Ap	629	16	0.7	273	10	US-09-878-574-12212	Sequence 12212, A
C 557	17	0.8	3422	9	US-09-764-891-6046	Sequence 6046, Ap	630	16	0.7	273	10	US-09-563-817-423	Sequence 423, App
C 558	17	0.8	3422	9	US-09-764-891-6046	Sequence 6046, Ap	631	16	0.7	277	10	US-09-878-574-15242	Sequence 15242, A
C 559	17	0.8	3422	9	US-10-091-438-271	Sequence 271, App	632	16	0.7	283	10	US-09-783-590-5853	Sequence 5853, Ap
C 560	17	0.8	3422	9	US-10-091-438-273	Sequence 273, App	633	16	0.7	285	10	US-09-294-093B-831	Sequence 831, App
C 561	17	0.8	3696	9	US-09-972-473-20	Sequence 20, Appl	634	16	0.7	286	10	US-09-878-574-14644	Sequence 14644, A
C 562	17	0.8	3718	10	US-09-917-800A-1471	Sequence 1471, Ap	635	16	0.7	304	10	US-09-864-761-21581	Sequence 21581, A
C 563	17	0.8	4086	9	US-10-161-803-18	Sequence 18, Appl	636	16	0.7	327	10	US-09-815-242-2750	Sequence 2750, Ap
C 564	17	0.8	4189	9	US-10-161-803-19	Sequence 19, Appl	637	16	0.7	338	10	US-09-867-701-7627	Sequence 7627, Ap
C 565	17	0.8	4315	12	US-10-044-090-526	Sequence 526, App	638	16	0.7	343	10	US-09-983-966-2317	Sequence 2317, Ap
C 566	17	0.8	4359	10	US-09-764-877-2346	Sequence 2346, Ap	639	16	0.7	345	10	US-09-878-574-2619	Sequence 2619, App
C 567	17	0.8	4360	10	US-09-764-877-2347	Sequence 2347, Ap	640	16	0.7	350	10	US-09-745-288-19	Sequence 19, Appl
C 568	17	0.8	4465	9	US-09-953-067A-2	Sequence 2, Appl	641	16	0.7	351	7	US-08-781-986A-2969	Sequence 2969, Ap
C 569	17	0.8	4466	9	US-10-071-766-46	Sequence 46, Appl	642	16	0.7	362	10	US-09-878-574-1744	Sequence 1744, Ap
C 570	17	0.8	4466	12	US-10-044-090-478	Sequence 478, App	643	16	0.7	364	9	US-10-015-219-665	Sequence 665, App
C 571	17	0.8	4998	10	US-09-738-968-42	Sequence 42, Appl	644	16	0.7	364	10	US-09-777-564-665	Sequence 665, App
C 572	17	0.8	5881	9	US-09-764-891-9918	Sequence 9918, Ap	645	16	0.7	366	9	US-10-015-219-792	Sequence 792, App
C 573	17	0.8	6226	10	US-09-827-949-1	Sequence 1, Appl	646	16	0.7	366	10	US-09-777-564-792	Sequence 792, App
C 574	17	0.8	7214	9	US-10-205-942-1	Sequence 2, Appl	647	16	0.7	367	10	US-09-925-301-78	Sequence 78, Appl
C 575	17	0.8	9453	9	US-09-954-456-324	Sequence 324, App	648	16	0.7	371	10	US-09-864-761-11731	Sequence 11731, A
C 576	17	0.8	10198	9	US-09-764-891-8743	Sequence 8743, Ap	649	16	0.7	372	9	US-09-918-995-37446	Sequence 37446, A
C 577	17	0.8	10198	9	US-09-764-891-8743	Sequence 8744, Ap	650	16	0.7	373	10	US-09-764-877-315	Sequence 315, App
C 578	17	0.8	11064	9	US-09-949-293-25	Sequence 25, Appl	651	16	0.7	377	9	US-09-918-995-29711	Sequence 29711, A
C 579	17	0.8	11735	9	US-10-239-676-34	Sequence 34, Appl	652	16	0.7	378	10	US-09-878-574-3365	Sequence 3365, Ap
C 580	17	0.8	14800	10	US-09-954-456-1601	Sequence 1601, Ap	653	16	0.7	380	9	US-09-918-995-7788	Sequence 7788, Ap
C 581	17	0.8	18404	9	US-10-142-515-10	Sequence 10, Appl	654	16	0.7	381	9	US-09-796-692-8999	Sequence 8999, Ap
C 582	17	0.8	31129	9	US-10-298-192-3	Sequence 3, Appl	655	16	0.7	381	9	US-10-040-862-8999	Sequence 8999, Ap
C 583	17	0.8	32154	10	US-09-764-877-3443	Sequence 3443, Ap	656	16	0.7	383	9	US-09-803-719-1392	Sequence 1392, Ap
C 584	17	0.8	32768	10	US-09-070-927A-17	Sequence 17, Appl	657	16	0.7	390	9	US-09-938-842A-2688	Sequence 2688, Ap
C 585	17	0.8	42571	9	US-10-224-413-3	Sequence 3, Appl	658	16	0.7	393	10	US-09-867-701-3338	Sequence 3238, Ap
C 586	17	0.8	53226	10	US-09-818-264-3	Sequence 3, Appl	659	16	0.7	393	10	US-09-974-300-7553	Sequence 7553, Ap
C 587	17	0.8	82326	9	US-09-818-264-3	Sequence 3, Appl	660	16	0.7	394	10	US-09-833-381-2021	Sequence 2021, Ap
C 588	17	0.8	180216	10	US-09-835-232-6	Sequence 6, Appl	661	16	0.7	395	9	US-10-060-036-300	Sequence 300, App
C 589	17	0.8	326014	10	US-09-731-231A-3	Sequence 3, Appl	662	16	0.7	397	10	US-09-960-352-10508	Sequence 10508, A
C 590	17	0.8	397658	10	US-09-813-320-3	Sequence 3, Appl	663	16	0.7	401	9	US-09-946-807-624	Sequence 624, App
C 591	17	0.8	513509	9	US-09-754-853A-4	Sequence 4, Appl	664	16	0.7	401	9	US-09-946-807-1125	Sequence 1125, Ap
C 592	17	0.8	513509	9	US-09-754-853A-4	Sequence 4, Appl	665	16	0.7	401	10	US-09-795-666-624	Sequence 624, App
C 593	17	0.8	1503841	9	US-09-846-807-1	Sequence 1, Appl	666	16	0.7	401	10	US-09-795-666-1125	Sequence 1125, App
C 594	17	0.8	1503841	10	US-09-795-668-1	Sequence 1, Appl	667	16	0.7	401	10	US-09-795-668-624	Sequence 624, App
C 595	17	0.8	1503841	10	US-09-795-668-1	Sequence 1, Appl	668	16	0.7	401	10	US-09-795-668-1125	Sequence 1125, Ap
C 596	17	0.8	1830121	9	US-09-738-626-1	Sequence 1, Appl	669	16	0.7	404	9	US-09-918-995-36717	Sequence 36717, A
C 597	17	0.8	3309400	9	US-09-738-626-1	Sequence 1, Appl	670	16	0.7	405	10	US-09-960-352-6748	Sequence 6748, Ap
C 598	17	0.8	3309400	9	US-09-738-626-1	Sequence 1, Appl	671	16	0.7	407	10	US-09-878-574-1001	Sequence 1001, Ap
C 599	17	0.8	3309400	9	US-09-738-626-1	Sequence 1, Appl	672	16	0.7	408	10	US-09-878-574-5018	Sequence 5018, Ap
C 600	16	0.7	112	10	US-09-864-761-28203	Sequence 28203, A	673	16	0.7	407	10	US-09-878-574-439	Sequence 439, App
C 601	16	0.7	118	9	US-10-005-338B-117	Sequence 117, App	674	16	0.7	409	10	US-09-880-107-971	Sequence 971, App
C 602	16	0.7	141	10	US-09-864-761-25877	Sequence 25877, A	675	16	0.7	421	9	US-09-918-995-6945	Sequence 6945, Ap
C 603	16	0.7	155	10	US-09-864-761-22570	Sequence 22570, A	676	16	0.7	421	9	US-09-918-995-36017	Sequence 36017, A

677	16	0.7	424	10	US-09-560-863-249	Sequence 249, App	c 750	16	0.7	702	9	US-10-091-504-2231	Sequence 2231, Ap
678	16	0.7	429	10	US-09-998-598-279	Sequence 279, App	c 751	16	0.7	702	9	US-10-091-504-2232	Sequence 2232, Ap
679	16	0.7	438	10	US-09-860-352-8396	Sequence 8396, Ap	c 752	16	0.7	702	10	US-09-764-865-2231	Sequence 2231, Ap
c 680	16	0.7	441	10	US-09-860-352-3512	Sequence 3512, Ap	c 753	16	0.7	702	10	US-09-764-865-2232	Sequence 2232, Ap
681	16	0.7	443	10	US-09-867-701-2506	Sequence 2506, App	c 754	16	0.7	703	10	US-09-910-943-302	Sequence 302, App
682	16	0.7	446	10	US-09-879-536-629	Sequence 629, App	c 755	16	0.7	704	9	US-10-091-504-2233	Sequence 2233, Ap
683	16	0.7	448	9	US-09-918-995-9655	Sequence 9655, App	c 756	16	0.7	704	10	US-09-764-865-2233	Sequence 2233, Ap
684	16	0.7	449	10	US-09-983-965-2802	Sequence 2802, App	c 757	16	0.7	714	9	US-10-144-929-72	Sequence 72, App1
685	16	0.7	455	10	US-09-867-550-61	Sequence 61, App1	c 758	16	0.7	750	9	US-10-101-464A-242	Sequence 242, App
686	16	0.7	455	10	US-09-867-701-3626	Sequence 3626, App	c 759	16	0.7	755	9	US-09-992-598-15	Sequence 15, App1
687	16	0.7	465	9	US-09-918-995-14180	Sequence 14180, A	c 760	16	0.7	755	9	US-09-889-293A-15	Sequence 15, App1
688	16	0.7	467	10	US-09-864-761-6179	Sequence 6179, App	c 761	16	0.7	755	9	US-09-889-735-15	Sequence 15, App1
689	16	0.7	468	10	US-09-864-761-15621	Sequence 15621, A	c 762	16	0.7	755	9	US-09-990-444-15	Sequence 15, App1
c 690	16	0.7	472	10	US-09-864-761-14629	Sequence 14629, A	c 763	16	0.7	755	9	US-09-989-730-15	Sequence 15, App1
c 691	16	0.7	472	10	US-09-764-877-726	Sequence 726, App	c 764	16	0.7	755	9	US-09-990-436-15	Sequence 15, App1
c 692	16	0.7	474	10	US-09-864-761-11618	Sequence 11618, A	c 765	16	0.7	755	9	US-09-991-181-15	Sequence 15, App1
c 693	16	0.7	476	7	US-08-781-986A-1469	Sequence 1469, App	c 766	16	0.7	755	9	US-09-993-687-15	Sequence 15, App1
c 694	16	0.7	477	9	US-09-918-995-9777	Sequence 9777, App	c 767	16	0.7	755	9	US-09-989-734-15	Sequence 15, App1
c 695	16	0.7	477	9	US-09-918-995-12297	Sequence 12297, A	c 768	16	0.7	755	9	US-09-997-653-15	Sequence 15, App1
c 696	16	0.7	478	9	US-09-998-598-654	Sequence 654, App	c 769	16	0.7	755	9	US-09-993-667-15	Sequence 15, App1
c 697	16	0.7	478	9	US-09-918-995-7269	Sequence 7269, App	c 770	16	0.7	755	9	US-09-990-438-15	Sequence 15, App1
c 698	16	0.7	479	9	US-09-918-995-15584	Sequence 35584, A	c 771	16	0.7	755	9	US-09-990-562-15	Sequence 15, App1
c 699	16	0.7	484	9	US-09-796-692-7	Sequence 7, App11	c 772	16	0.7	755	9	US-09-997-428-15	Sequence 15, App1
700	16	0.7	484	9	US-09-796-692-3311	Sequence 3311, App	c 773	16	0.7	755	9	US-09-997-666-15	Sequence 15, App1
701	16	0.7	484	9	US-09-918-995-19639	Sequence 19639, A	c 774	16	0.7	755	9	US-09-990-711-15	Sequence 15, App1
702	16	0.7	484	9	US-10-040-862-7	Sequence 7, App11	c 775	16	0.7	755	9	US-09-989-726-15	Sequence 15, App1
703	16	0.7	484	9	US-10-040-862-3311	Sequence 3311, App	c 776	16	0.7	755	9	US-09-990-437-15	Sequence 15, App1
704	16	0.7	487	9	US-10-015-219-1218	Sequence 1218, App	c 777	16	0.7	755	9	US-09-998-156-15	Sequence 15, App1
705	16	0.7	487	9	US-09-777-564-1218	Sequence 1218, App	c 778	16	0.7	755	9	US-09-991-157-15	Sequence 15, App1
c 706	16	0.7	489	10	US-10-001-883-30	Sequence 30, App1	c 779	16	0.7	755	9	US-09-991-172-15	Sequence 15, App1
707	16	0.7	491	9	US-09-918-995-20779	Sequence 20779, A	c 780	16	0.7	755	9	US-09-997-514-15	Sequence 15, App1
708	16	0.7	495	9	US-09-918-995-20203	Sequence 20203, A	c 781	16	0.7	755	9	US-09-997-573-15	Sequence 15, App1
709	16	0.7	495	9	US-09-918-995-21514	Sequence 21514, A	c 782	16	0.7	755	9	US-09-990-443-15	Sequence 15, App1
710	16	0.7	495	10	US-09-880-107-1140	Sequence 1140, App	c 783	16	0.7	755	9	US-09-990-726-15	Sequence 15, App1
711	16	0.7	496	9	US-10-046-935-407	Sequence 407, App	c 784	16	0.7	755	9	US-09-997-559-15	Sequence 15, App1
712	16	0.7	496	9	US-09-878-178-407	Sequence 407, App	c 785	16	0.7	755	9	US-09-997-601-15	Sequence 15, App1
713	16	0.7	496	9	US-10-146-502-407	Sequence 407, App	c 786	16	0.7	755	9	US-09-989-729A-15	Sequence 15, App1
c 714	16	0.7	497	10	US-09-864-761-10862	Sequence 10862, A	c 787	16	0.7	755	9	US-09-990-440-15	Sequence 15, App1
715	16	0.7	501	9	US-09-938-842A-2038	Sequence 2038, App	c 788	16	0.7	755	9	US-09-991-854-15	Sequence 15, App1
716	16	0.7	501	9	US-09-954-692-97	Sequence 97, App1	c 789	16	0.7	755	9	US-09-997-349-15	Sequence 15, App1
717	16	0.7	501	10	US-09-559-671A-97	Sequence 97, App1	c 790	16	0.7	755	9	US-09-997-440-15	Sequence 15, App1
c 718	16	0.7	502	10	US-09-864-761-13439	Sequence 13439, A	c 791	16	0.7	755	9	US-09-997-628-15	Sequence 15, App1
719	16	0.7	505	9	US-10-091-504-1792	Sequence 1792, App	c 792	16	0.7	755	9	US-09-997-683-15	Sequence 15, App1
720	16	0.7	505	10	US-09-764-869-1792	Sequence 1792, App	c 793	16	0.7	755	9	US-09-993-469-15	Sequence 15, App1
c 721	16	0.7	508	9	US-10-060-036-1137	Sequence 1137, App	c 794	16	0.7	755	9	US-09-993-748-15	Sequence 15, App1
c 722	16	0.7	509	10	US-09-998-598-522	Sequence 522, App	c 795	16	0.7	755	9	US-09-997-542-15	Sequence 15, App1
723	16	0.7	510	9	US-09-918-995-32343	Sequence 32343, A	c 796	16	0.7	755	9	US-09-990-427-15	Sequence 15, App1
c 724	16	0.7	516	9	US-10-015-219-944	Sequence 944, App	c 797	16	0.7	755	9	US-09-990-439-15	Sequence 15, App1
725	16	0.7	516	10	US-09-777-564-944	Sequence 944, App	c 798	16	0.7	755	9	US-09-989-328-15	Sequence 15, App1
726	16	0.7	517	10	US-09-974-300-3195	Sequence 3195, App	c 799	16	0.7	755	9	US-09-993-583-15	Sequence 15, App1
c 727	16	0.7	518	10	US-09-864-761-7886	Sequence 7886, App	c 800	16	0.7	755	9	US-09-941-992-15	Sequence 15, App1
c 728	16	0.7	527	10	US-09-833-381-1083	Sequence 1083, App	c 801	16	0.7	755	9	US-09-992-521-15	Sequence 15, App1
c 729	16	0.7	530	10	US-09-867-701-5223	Sequence 5223, App	c 802	16	0.7	755	9	US-09-997-333-15	Sequence 15, App1
730	16	0.7	533	10	US-09-867-701-5295	Sequence 5295, App	c 803	16	0.7	755	9	US-09-997-384-15	Sequence 15, App1
731	16	0.7	541	9	US-10-091-504-1793	Sequence 1793, App	c 804	16	0.7	755	10	US-09-989-722-15	Sequence 15, App1
732	16	0.7	541	9	US-09-764-869-1793	Sequence 1793, App	c 805	16	0.7	755	10	US-09-989-723-15	Sequence 15, App1
c 733	16	0.7	547	9	US-09-991-936-101	Sequence 301, App	c 806	16	0.7	755	10	US-09-989-727-15	Sequence 15, App1
c 734	16	0.7	549	10	US-09-864-761-12848	Sequence 12848, A	c 807	16	0.7	755	10	US-09-989-727-15	Sequence 15, App1
735	16	0.7	558	10	US-09-998-598-586	Sequence 586, App	c 808	16	0.7	755	10	US-09-989-727-15	Sequence 15, App1
736	16	0.7	565	10	US-09-827-669-56	Sequence 56, App1	c 809	16	0.7	755	10	US-09-989-733-15	Sequence 15, App1
c 737	16	0.7	566	10	US-09-864-761-9834	Sequence 9834, App	c 810	16	0.7	755	10	US-09-991-073-15	Sequence 15, App1
c 738	16	0.7	573	10	US-09-998-598-1184	Sequence 1184, App	c 811	16	0.7	755	10	US-09-990-442-15	Sequence 15, App1
739	16	0.7	587	10	US-09-864-761-9344	Sequence 9344, App	c 812	16	0.7	755	10	US-09-991-163-15	Sequence 15, App1
c 740	16	0.7	589	9	US-10-015-219-101	Sequence 101, App	c 813	16	0.7	755	10	US-09-993-606-15	Sequence 15, App1
c 741	16	0.7	589	10	US-09-777-564-101	Sequence 101, App	c 814	16	0.7	755	10	US-09-990-456-15	Sequence 15, App1
742	16	0.7	602	10	US-09-998-598-1275	Sequence 1275, App	c 815	16	0.7	755	10	US-09-989-721-15	Sequence 15, App1
c 743	16	0.7	621	9	US-09-796-692-3953	Sequence 3953, App	c 816	16	0.7	755	10	US-09-991-073-15	Sequence 15, App1
c 744	16	0.7	621	9	US-10-040-862-3953	Sequence 3953, App	c 817	16	0.7	755	10	US-09-815-244-4379	Sequence 4379, App
745	16	0.7	625	10	US-09-770-149-748	Sequence 748, App	c 818	16	0.7	765	9	US-10-066-543-119	Sequence 119, App
746	16	0.7	659	10	US-09-881-752A-277	Sequence 277, App	c 819	16	0.7	765	10	US-09-815-244-8176	Sequence 8176, App
c 747	16	0.7	661	10	US-09-822-849A-584	Sequence 584, App	c 820	16	0.7	782	9	US-09-809-391-283	Sequence 283, App
748	16	0.7	666	9	US-10-198-846-8598	Sequence 8598, App	c 821	16	0.7	804	10	US-09-770-445-802	Sequence 802, App
749	16	0.7	694	10	US-09-815-242-4776	Sequence 4776, App	c 822	16	0.7	833	10	US-09-770-445-686	Sequence 686, App

823	16	0.7	833	12	US-10-044-090-139	Sequence 339, App	896	16	0.7	906	9	US-10-176-751-541	Sequence 541, App
824	16	0.7	846	9	US-10-001-835-96	Sequence 96, App	897	16	0.7	906	9	US-10-176-760-541	Sequence 541, App
825	16	0.7	847	9	US-10-198-846-1684	Sequence 1684, Ap	898	16	0.7	906	9	US-10-176-990-541	Sequence 541, App
826	16	0.7	870	10	US-09-974-300-6506	Sequence 6506, Ap	899	16	0.7	906	9	US-10-180-554-541	Sequence 541, App
827	16	0.7	906	9	US-10-174-590-541	Sequence 541, App	900	16	0.7	906	9	US-10-180-544-541	Sequence 541, App
828	16	0.7	906	9	US-10-176-758-541	Sequence 541, App	901	16	0.7	906	9	US-10-180-548-541	Sequence 541, App
829	16	0.7	906	9	US-10-175-737-541	Sequence 541, App	902	16	0.7	906	9	US-10-180-551-541	Sequence 541, App
830	16	0.7	906	9	US-10-173-706-541	Sequence 541, App	903	16	0.7	906	9	US-10-180-998-541	Sequence 541, App
831	16	0.7	906	9	US-10-175-738-541	Sequence 541, App	904	16	0.7	906	9	US-10-180-999-541	Sequence 541, App
832	16	0.7	906	9	US-10-175-752-541	Sequence 541, App	905	16	0.7	906	9	US-10-183-011-541	Sequence 541, App
833	16	0.7	906	9	US-10-176-482-541	Sequence 541, App	906	16	0.7	906	9	US-10-184-612-541	Sequence 541, App
834	16	0.7	906	9	US-10-176-787-541	Sequence 541, App	907	16	0.7	906	9	US-10-184-615-541	Sequence 541, App
835	16	0.7	906	9	US-10-176-913-541	Sequence 541, App	908	16	0.7	906	9	US-10-184-617-541	Sequence 541, App
836	16	0.7	906	9	US-10-180-552-541	Sequence 541, App	909	16	0.7	906	9	US-10-184-622-541	Sequence 541, App
837	16	0.7	906	9	US-10-180-557-541	Sequence 541, App	910	16	0.7	906	9	US-10-184-628-541	Sequence 541, App
838	16	0.7	906	9	US-10-173-700-541	Sequence 541, App	911	16	0.7	906	9	US-10-184-632-541	Sequence 541, App
839	16	0.7	906	9	US-10-174-572-541	Sequence 541, App	912	16	0.7	906	9	US-10-184-635-541	Sequence 541, App
840	16	0.7	906	9	US-10-174-579-541	Sequence 541, App	913	16	0.7	906	9	US-10-184-637-541	Sequence 541, App
841	16	0.7	906	9	US-10-174-582-541	Sequence 541, App	914	16	0.7	906	9	US-10-184-638-541	Sequence 541, App
842	16	0.7	906	9	US-10-174-588-541	Sequence 541, App	915	16	0.7	906	9	US-10-184-640-541	Sequence 541, App
843	16	0.7	906	9	US-10-175-739-541	Sequence 541, App	916	16	0.7	906	9	US-10-184-646-541	Sequence 541, App
844	16	0.7	906	9	US-10-175-740-541	Sequence 541, App	917	16	0.7	906	9	US-10-184-650-541	Sequence 541, App
845	16	0.7	906	9	US-10-175-740-541	Sequence 541, App	918	16	0.7	906	9	US-10-184-653-541	Sequence 541, App
846	16	0.7	906	9	US-10-176-488-541	Sequence 541, App	919	16	0.7	906	9	US-10-187-588-541	Sequence 541, App
847	16	0.7	906	9	US-10-176-492-541	Sequence 541, App	920	16	0.7	906	9	US-10-187-597-541	Sequence 541, App
848	16	0.7	906	9	US-10-176-747-541	Sequence 541, App	921	16	0.7	906	9	US-10-187-598-541	Sequence 541, App
849	16	0.7	906	9	US-10-176-750-541	Sequence 541, App	922	16	0.7	906	9	US-10-187-600-541	Sequence 541, App
850	16	0.7	906	9	US-10-176-985-541	Sequence 541, App	923	16	0.7	906	9	US-10-187-601-541	Sequence 541, App
851	16	0.7	906	9	US-10-176-987-541</								

```
c 969 16 0.7 906 9 US-10-195-902-541 Sequence 541, App
c 970 16 0.7 906 9 US-10-196-743-541 Sequence 541, App
c 971 16 0.7 906 9 US-10-196-745-541 Sequence 541, App
c 972 16 0.7 906 9 US-10-196-760-541 Sequence 541, App
c 973 16 0.7 906 9 US-10-196-762-541 Sequence 541, App
c 974 16 0.7 906 9 US-10-197-695-541 Sequence 541, App
c 975 16 0.7 906 9 US-10-176-484-541 Sequence 541, App
c 976 16 0.7 906 9 US-10-176-753-541 Sequence 541, App
c 977 16 0.7 906 9 US-10-176-917-541 Sequence 541, App
c 978 16 0.7 906 9 US-10-176-982-541 Sequence 541, App
c 979 16 0.7 906 9 US-10-179-506-541 Sequence 541, App
c 980 16 0.7 906 9 US-10-179-513-541 Sequence 541, App
c 981 16 0.7 906 9 US-10-179-514-541 Sequence 541, App
c 982 16 0.7 906 9 US-10-179-522-541 Sequence 541, App
c 983 16 0.7 906 9 US-10-180-556-541 Sequence 541, App
c 984 16 0.7 906 9 US-10-180-560-541 Sequence 541, App
c 985 16 0.7 906 9 US-10-183-015-541 Sequence 541, App
c 986 16 0.7 906 9 US-10-184-615-541 Sequence 541, App
c 987 16 0.7 906 9 US-10-184-620-541 Sequence 541, App
c 988 16 0.7 906 9 US-10-184-643-541 Sequence 541, App
c 989 16 0.7 906 9 US-10-184-656-541 Sequence 541, App
c 990 16 0.7 906 9 US-10-192-010-541 Sequence 541, App
c 991 16 0.7 906 9 US-10-195-894-541 Sequence 541, App
c 992 16 0.7 906 9 US-10-205-908-541 Sequence 541, App
c 993 16 0.7 906 9 US-10-186-855-541 Sequence 541, App
c 994 16 0.7 906 9 US-10-187-599-541 Sequence 541, App
c 995 16 0.7 906 9 US-10-187-750-541 Sequence 541, App
c 996 16 0.7 906 9 US-10-188-780-541 Sequence 541, App
c 997 16 0.7 906 9 US-10-192-015-541 Sequence 541, App
c 998 16 0.7 906 9 US-10-194-394-541 Sequence 541, App
c 999 16 0.7 906 9 US-10-194-425-541 Sequence 541, App
c1000 16 0.7 906 9 US-10-194-485-541 Sequence 541, App
```

## ALIGNMENTS

```
RESULT 1
US-09-833-381-1475
; Sequence 1475, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1475
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-833-381-1475
```

```
Query Match 15.7%; Score 340; DB 10; Length 420;
Best Local Similarity 99.7%; Pred. No. 1e-177;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 807 CAATACGACATATTTCTGTGAGAGTGTGACATGATGACACACCAATTCGATGAC 866
DB 30 CAATACGACATATTTCTGTGAGAGTGTGACATGATGACACACCAATTCGATGAC 89
QY 867 TTTTGAAGAAGCTAATCTCTGTGACACCTGAATTAACACATTTGCTTAATAC 926
DB 90 TTTTGAAGAAGCTAATCTCTGTGACACCTGAATTAACACATTTGCTTAATAC 149
QY 927 TAAGCTTACCTCGTGGAAAAATACGATATTCCTATCATCTTGGAGGAGAGATTTGAT 986
DB 150 TAAGCTTACCTCGTGGAAAAATACGATATTCCTATCATCTTGGAGGAGAGATTTGAT 209
```

```
QY 987 GGCTTGTGCTCAACAGGCTGTGGAGAGTGGCGCTTTCCTACCAATTTGGCTCA 1046
DB 210 GGCTTGTGCTCAACAGGCTGTGGAGAGTGGCGCTTTCCTACCAATTTGGCTCA 269
QY 1047 TATGATGCTATGATGATATTAATCTGCTGCTTTTAAGAGTTGACGAACACAGTGTAT 1106
DB 270 TATGATGCTATGATGATATTAATCTGCTGCTTTTAAGAGTTGACGAACACAGTGTAT 329
QY 1107 TATTTGAGACCAACATCGAATTTGGTCAACAGATTTATTTTGAAGCCAGAAATTTTC 1166
DB 330 TATTTGAGACCAACATCGAATTTGGTCAACAGATTTATTTTGAAGCCAGAAATTTTC 389
QY 1167 TTTTGGACTTGTGTAAAGCTGTGTTATTA 1197
DB 390 TTTTGGACTTGTGTAAAGCTGTGTTATTA 420
```

```
RESULT 2
US-09-764-891-1170
; Sequence 1170, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1170
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-891-1170
```

```
Query Match 11.7%; Score 255; DB 9; Length 591;
Best Local Similarity 100.0%; Pred. No. 1.6e-130;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1883 TTGGGGCTACTGCTGCTGTGTGGGAATATCGCAGAGCAATTTCTTTTGTATCTTGAAAT 1942
DB 30 TTGGGGCTACTGCTGCTGTGTGGGAATATCGCAGAGCAATTTCTTTTGTATCTTGAAAT 89
QY 1943 CGGATACCATTTAGACAGACCTCTAGTAAGATTTGACAGATGCTCAACAGATGTTG 2002
DB 90 CGGATACCATTTAGACAGACCTCTAGTAAGATTTGACAGATGCTCAACAGATGTTG 149
QY 2003 CTGCATGTTGGAAGAAATGCTTTAGTACATACATTCCTGCTTCACTGGTAGTACAA 2062
DB 150 CTGCATGTTGGAAGAAATGCTTTAGTACATACATTCCTGCTTCACTGGTAGTACAA 209
QY 2063 GAGGAAAGCTGTTTGTATAGTTGATACCAAGGAGGCAAGCACTTTGAACACAGCTG 2122
DB 210 GAGGAAAGCTGTTTGTATAGTTGATACCAAGGAGGCAAGCACTTTGAACACAGCTG 269
QY 2123 GGTTCCTCTTCTAC 2137
DB 270 GGTTCCTCTTCTAC 284
```

```
RESULT 3
US-09-764-891-7481
; Sequence 7481, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
```



SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 7481  
LENGTH: 2194  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-891-7481

Query Match 10.1%; Score 220; DB 9; Length 2194;  
Best Local Similarity 100.0%; Pred. No. 4.9e-111;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1768 AAGTGGCCAGTTCTTGTGTTGCTACTGATGCTGCGAGAGGGCTGGATATTGAAATGTG 1827  
DB 1 AAGTGGCCAGTTCTTGTGTTGCTACTGATGCTGCGAGAGGGCTGGATATTGAAATGTG 60  
QY 1828 CAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGATGTTTCAATGGAATGGG 1887  
DB 61 CAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGATGTTTCAATGGAATGGG 120  
QY 1888 CGTACTGTCGTTGTGGGAATAGTGGCAGAGCAATTCCTTTTGAATCTTGAATCGGAT 1947  
DB 121 CGTACTGTCGTTGTGGGAATAGTGGCAGAGCAATTCCTTTTGAATCTTGAATCGGAT 180  
QY 1948 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 1987  
DB 181 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 220

RESULT 4  
US-09-764-891-7482  
Sequence 7482, Application US/09764891  
Publication No. US20030077808A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC006  
CURRENT APPLICATION NUMBER: US/09/764,891  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 10231  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 7482  
LENGTH: 2195  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-891-7482

Query Match 10.1%; Score 220; DB 9; Length 2195;  
Best Local Similarity 100.0%; Pred. No. 4.9e-111;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1768 AAGTGGCCAGTTCTTGTGTTGCTACTGATGCTGCGAGAGGGCTGGATATTGAAATGTG 1827  
DB 1 AAGTGGCCAGTTCTTGTGTTGCTACTGATGCTGCGAGAGGGCTGGATATTGAAATGTG 60  
QY 1828 CAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGATGTTTCAATGGAATGGG 1887  
DB 61 CAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGATGTTTCAATGGAATGGG 120  
QY 1888 CGTACTGTCGTTGTGGGAATAGTGGCAGAGCAATTCCTTTTGAATCTTGAATCGGAT 1947  
DB 121 CGTACTGTCGTTGTGGGAATAGTGGCAGAGCAATTCCTTTTGAATCTTGAATCGGAT 180  
QY 1948 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 1987  
DB 181 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 220

RESULT 5  
US-09-764-891-7483  
Sequence 7483, Application US/09764891  
Publication No. US20030077808A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC006  
CURRENT APPLICATION NUMBER: US/09/764,891  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 10231  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 7483  
LENGTH: 2197  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-891-7483

Query Match 10.1%; Score 220; DB 9; Length 2197;  
Best Local Similarity 100.0%; Pred. No. 4.9e-111;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1768 AAGTGGCCAGTTCTTGTGTTGCTACTGATGCTGCGAGAGGGCTGGATATTGAAATGTG 1827  
DB 1 AAGTGGCCAGTTCTTGTGTTGCTACTGATGCTGCGAGAGGGCTGGATATTGAAATGTG 60  
QY 1828 CAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGATGTTTCAATGGAATGGG 1887  
DB 61 CAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGATGTTTCAATGGAATGGG 120  
QY 1888 CGTACTGTCGTTGTGGGAATAGTGGCAGAGCAATTCCTTTTGAATCTTGAATCGGAT 1947  
DB 121 CGTACTGTCGTTGTGGGAATAGTGGCAGAGCAATTCCTTTTGAATCTTGAATCGGAT 180  
QY 1948 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 1987  
DB 181 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 220

RESULT 6  
US-09-815-242-6893

Sequence 6893, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
Prior Application Number: 60/191,078  
Prior Filing Date: 2000-03-21  
Prior Application Number: 60/206,848  
Prior Filing Date: 2000-05-23  
Prior Application Number: 60/207,727  
Prior Filing Date: 2000-05-26  
Prior Application Number: 60/242,578  
Prior Filing Date: 2000-10-23  
Prior Application Number: 60/253,625  
Prior Filing Date: 2000-11-27  
Prior Application Number: 60/257,931  
Prior Filing Date: 2000-12-22  
Prior Application Number: 60/269,308  
Prior Filing Date: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6893  
LENGTH: 1842  
TYPE: DNA  
ORGANISM: Haemophilus influenzae

FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1842)  
US-09-815-242-6893

Query Match 1.0%; Score 21; DB 10; Length 1842;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1879 CGAATGGGCTACTGCTGCT 1899  
Db 991 CGAATGGGCTACTGCTGCT 1011

RESULT 7  
US-10-329-960-1  
Sequence 1, Application US/10329960  
Publication No. US20030099277A1  
GENERAL INFORMATION:  
APPLICANT: Fleischmann et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag  
FILE REFERENCE: PB186P1  
CURRENT APPLICATION NUMBER: US/10/329,960  
PRIOR APPLICATION NUMBER: US 09/643,990  
PRIOR FILING DATE: 2003-01-02  
PRIOR APPLICATION NUMBER: US 08/487,429  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/426,787  
PRIOR FILING DATE: 1995-04-21  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1830121  
TYPE: DNA  
ORGANISM: Haemophilus influenzae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4747)..(4747)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (9921)..(9921)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (10150)..(10150)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (29298)..(29298)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36543)..(36543)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36551)..(36551)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (36636)..(36636)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (40808)..(40810)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44416)..(44416)  
OTHER INFORMATION: n equals a, t, g or c

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44905)..(44905)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (44975)..(44975)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45593)..(45593)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (45732)..(45732)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (47036)..(47036)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51334)..(51334)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51786)..(51786)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (51805)..(51805)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (55369)..(55369)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (65309)..(65309)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (65313)..(65313)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (80024)..(80024)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (100091)..(100091)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (102696)..(102696)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (105121)..(105121)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (107248)..(107248)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (117136)..(117136)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (119750) ..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924) ..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038) ..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344) ..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167) ..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336) ..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340) ..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) ..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) ..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) ..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) ..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) ..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) ..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) ..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) ..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) ..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) ..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530) ..(152530)
```

Query Match 1.0%; Score 21; DB 9; Length 1830121;  
Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1879 CGAATTGGCGTACTGTCGT 1899  
Db 261968 CGAATTGGCGTACTGTCGT 261988

RESULT 8  
US-09-878-574-10286  
Sequence 10286, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 10286  
LENGTH: 244  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: 701103507H1  
US-09-878-574-10286

Query Match 0.9%; Score 20; DB 10; Length 244;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 686 CAGAAGCAGAGAGAGAGAA 705  
Db 8 CAGAAGCAGAGAGAGAGAA 27

RESULT 9  
US-09-878-574-12453  
Sequence 12453, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 12453  
LENGTH: 266  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: 701065791H1  
US-09-878-574-12453

Query Match 0.9%; Score 20; DB 10; Length 266;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 686 CAGAAGCAGAGAGAGAGAA 705  
Db 3 CAGAAGCAGAGAGAGAGAA 22

RESULT 10  
US-09-864-761-28835/c

```
; Sequence 28835, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28835
; LENGTH: 138
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132775.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
; OTHER INFORMATION: EST HUMAN HIT: H69214.1, EVALUATE 2.00e-24
; OTHER INFORMATION: NT HIT: AL163283.2, EVALUATE 9.00e-14
; US-09-864-761-28835

Query Match      0.9%; Score 19; DB 10; Length 138;
Best Local Similarity 100.0%; Pred.No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 11
US-09-918-995-35717
; Sequence 35717, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35717
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(418)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-35717

Query Match      0.9%; Score 19; DB 9; Length 418;
Best Local Similarity 100.0%; Pred.No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1829 AACATGTTATCATTTTGA 1847
Db      249 AACATGTTATCATTTTGA 267

RESULT 12
US-09-864-761-12255/C
; Sequence 12255, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
```

;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 12255  
;; LENGTH: 592  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL32775.14  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.98  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.96  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87  
US-09-864-761-12255

Query Match 0.9%; Score 19; DB 10; Length 592;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 CAGAAGAGGTGGAAGAGT 453  
Db 320 CAGAAGAGGTGGAAGAGT 302

RESULT 13  
US-09-925-299-60/c  
;; Sequence 60, Application US/09925299  
;; Publication No. US20030040617A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
;; FILE REFERENCE: PA102  
;; CURRENT APPLICATION NUMBER: US/09/925,299  
;; PRIOR FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: PCT/US00/05883  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 1556  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 60  
;; LENGTH: 1472  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (129)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc\_feature  
;; LOCATION: (130)  
;; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-60

Query Match 0.9%; Score 19; DB 9; Length 1472;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 AGCTGTGTTCTGAGAG 334  
|||||

Db 354 AGCTGTGTTCTGAGAG 336

RESULT 14  
US-09-925-299-60/c  
;; Sequence 60, Application US/09925299  
;; Patent No. US20020055627A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
;; FILE REFERENCE: PA102  
;; CURRENT APPLICATION NUMBER: US/09/925,299  
;; PRIOR FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: PCT/US00/05883  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 1556  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 60  
;; LENGTH: 1472  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (129)  
;; OTHER INFORMATION: n equals a,t,g, or c  
;; NAME/KEY: misc\_feature  
;; LOCATION: (130)  
;; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-60

Query Match 0.9%; Score 19; DB 10; Length 1472;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 AGCTGTGTTCTGAGAG 334  
Db 354 AGCTGTGTTCTGAGAG 336

RESULT 15  
US-09-925-301-185  
;; Sequence 185, Application US/09925301  
;; Patent No. US20020052308A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
;; FILE REFERENCE: PA106  
;; CURRENT APPLICATION NUMBER: US/09/925,301  
;; PRIOR FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: PCT/US00/05882  
;; PRIOR FILING DATE: 2000-03-08  
;; PRIOR APPLICATION NUMBER: 60/124,270  
;; PRIOR FILING DATE: 1999-03-12  
;; NUMBER OF SEQ ID NOS: 1694  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 185  
;; LENGTH: 1987  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (523)  
;; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-301-185

Query Match 0.9%; Score 19; DB 10; Length 1987;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1434 TATGTTCACTGCACTTTT 1452  
|||||

Wed Jun 11 10:49:16 2003

us-09-714-865-15.oligo.rnpb

Page 14

Db 1074 TATGTCAGTGCACACTTT 1092

Search completed: June 10, 2003, 16:35:47  
Job time : 335 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 12:19:45 ; Search time 130 Seconds  
(without alignments)  
5123.861 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 2172  
Sequence: 1 atcggggagatgaagatctggga.....tagatgatgagtcagtcggat 2172

Scoring table: OLIGO\_NBC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued Patents: NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	0.9	2319	3	US-09-058-489-90
2	19	0.9	2850	4	US-09-184-964-2
3	19	0.9	3408	3	US-09-058-489-14
4	19	0.9	4416	3	US-09-058-489-17
5	19	0.9	5322	3	US-09-058-489-13
6	18	0.8	339	4	US-08-651-155B-116
7	18	0.8	421	4	US-09-222-575-116
8	18	0.8	532	4	US-09-318-443-3
9	18	0.8	1536	4	US-09-318-443-5
10	18	0.8	1682	4	US-09-318-443-7
11	18	0.8	1796	4	US-09-255-392-1
12	18	0.8	1876	4	US-09-082-092-5
13	18	0.8	1914	4	US-09-291-922-19
14	18	0.8	2005	2	US-08-484-200-1
15	18	0.8	2005	2	US-08-477-493-1
16	18	0.8	2005	3	US-08-465-375-1
17	18	0.8	2005	3	US-09-788-070-1
18	18	0.8	3103	3	US-08-826-246-3
19	18	0.8	3103	3	US-08-944-495-3
20	18	0.8	3103	3	US-09-126-640-2
21	18	0.8	3103	4	US-08-925-588-3
22	18	0.8	3103	4	US-09-288-292A-2
23	18	0.8	3111	3	US-09-487-444-3
24	18	0.8	11120	4	US-09-201-227A-7
25	17	0.8	624	4	US-09-161-241-6
26	17	0.8	780	4	US-09-161-241-5
27	17	0.8	838	3	US-09-062-416-19

C 28	17	0.8	1196	1	US-08-167-939A-5	Sequence 5, Appl
C 29	17	0.8	1196	1	US-08-567-538-5	Sequence 5, Appl
C 30	17	0.8	1237	1	US-08-197-792-40	Sequence 40, Appl
C 31	17	0.8	1237	1	US-08-459-850-40	Sequence 40, Appl
C 32	17	0.8	1237	1	US-08-459-214-40	Sequence 40, Appl
C 33	17	0.8	1254	5	PCT-US86-05320A-894	Sequence 894, App
C 34	17	0.8	1401	2	US-08-811-897A-32	Sequence 32, Appl
C 35	17	0.8	1401	2	US-08-855-213-32	Sequence 32, Appl
C 36	17	0.8	1401	4	US-09-201-474-32	Sequence 32, Appl
C 37	17	0.8	1446	4	US-09-537-357-8	Sequence 8, Appl
C 38	17	0.8	1485	2	US-08-811-897A-33	Sequence 33, Appl
C 39	17	0.8	1485	2	US-08-855-213-33	Sequence 33, Appl
C 40	17	0.8	1485	4	US-09-201-474-33	Sequence 33, Appl
C 41	17	0.8	1578	1	US-08-188-582-29	Sequence 29, Appl
C 42	17	0.8	1578	1	US-08-646-715-29	Sequence 29, Appl
C 43	17	0.8	1617	4	US-09-532-594B-19	Sequence 19, Appl
C 44	17	0.8	1715	4	US-09-537-357-56	Sequence 56, Appl
C 45	17	0.8	1800	4	US-09-532-594B-17	Sequence 17, Appl
C 46	17	0.8	1869	2	US-08-811-897A-40	Sequence 40, Appl
C 47	17	0.8	1869	2	US-08-855-213-40	Sequence 40, Appl
C 48	17	0.8	1869	4	US-09-201-474-40	Sequence 40, Appl
C 49	17	0.8	1924	3	US-08-961-083-159	Sequence 159, App
C 50	17	0.8	1929	2	US-08-746-283-4	Sequence 4, Appl
C 51	17	0.8	2179	4	US-08-858-207A-37	Sequence 37, Appl
C 52	17	0.8	2208	4	US-09-532-594B-5	Sequence 5, Appl
C 53	17	0.8	2386	2	US-08-746-283-2	Sequence 2, Appl
C 54	17	0.8	2386	2	US-08-746-257A-2	Sequence 2, Appl
C 55	17	0.8	2386	4	US-09-347-878A-47	Sequence 47, Appl
C 56	17	0.8	2548	2	US-08-811-897A-41	Sequence 41, Appl
C 57	17	0.8	2548	2	US-08-855-213-41	Sequence 41, Appl
C 58	17	0.8	2548	4	US-09-201-474-41	Sequence 41, Appl
C 59	17	0.8	2608	1	US-08-160-861-1	Sequence 1, Appl
C 60	17	0.8	2846	4	US-08-915-795-1	Sequence 1, Appl
C 61	17	0.8	3124	4	US-09-325-932A-111	Sequence 131, App
C 62	17	0.8	3293	4	US-09-104-324B-1	Sequence 1, Appl
C 63	17	0.8	3353	4	US-09-162-713-1	Sequence 1, Appl
C 64	17	0.8	3752	4	US-08-961-527-208	Sequence 208, App
C 65	17	0.8	42571	4	US-09-810-347-3	Sequence 3, Appl
C 66	17	0.8	4767	4	US-09-532-594A-1	Sequence 1, Appl
C 67	17	0.8	6226	4	US-08-542-363-1	Sequence 1, Appl
C 68	17	0.8	6226	4	US-09-100-089-1	Sequence 1, Appl
C 69	17	0.8	6226	4	US-09-670-827-1	Sequence 1, Appl
C 70	17	0.8	8512	4	US-08-961-527-129	Sequence 129, App
C 71	17	0.8	82571	4	US-09-810-347-3	Sequence 3, Appl
C 72	17	0.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
C 73	17	0.8	246240	2	US-08-724-394A-21	Sequence 21, Appl
C 74	17	0.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 75	17	0.8	4403765	4	US-09-103-840A-2	Sequence 2, Appl
C 76	16	0.7	180	3	US-09-103-840A-1	Sequence 1, Appl
C 77	16	0.7	160	3	US-08-466-548B-5	Sequence 5, Appl
C 78	16	0.7	160	3	US-07-998-226A-5	Sequence 5, Appl
C 79	16	0.7	178	1	US-08-248-474-62	Sequence 62, Appl
C 80	16	0.7	224	3	US-08-756-849-62	Sequence 62, Appl
C 81	16	0.7	224	4	US-09-397-787-105	Sequence 105, App
C 82	16	0.7	224	4	US-09-134-001C-1540	Sequence 1540, App
C 83	16	0.7	252	4	US-08-305-223-232	Sequence 232, App
C 84	16	0.7	252	4	US-09-149-476-142	Sequence 142, App
C 85	16	0.7	350	4	US-09-118-554-19	Sequence 19, Appl
C 86	16	0.7	350	4	US-09-118-627-19	Sequence 19, Appl
C 87	16	0.7	350	4	US-09-602-877A-19	Sequence 19, Appl
C 88	16	0.7	446	4	US-09-328-111-629	Sequence 629, App
C 89	16	0.7	501	4	US-09-339-913B-97	Sequence 97, Appl
C 90	16	0.7	501	4	US-09-339-904A-97	Sequence 97, Appl
C 91	16	0.7	501	4	US-08-769-062B-97	Sequence 97, Appl
C 92	16	0.7	501	4	US-09-344-002B-97	Sequence 97, Appl
C 93	16	0.7	501	4	US-09-559-565C-97	Sequence 97, Appl
C 94	16	0.7	555	4	US-09-712-016-56	Sequence 56, Appl
C 95	16	0.7	641	4	US-08-961-527-359	Sequence 359, App
C 96	16	0.7	702	4	US-09-134-001C-806	Sequence 806, App
C 97	16	0.7	7924	2	US-09-149-476-283	Sequence 283, App
C 98	16	0.7	934	2	US-08-867-030B-4	Sequence 4, Appl
C 99	16	0.7	934	5	PCT-US95-0611B-4	Sequence 4, Appl
C 100	16	0.7	1191	4	US-08-858-207A-124	Sequence 124, App
					US-09-181-336-18	Sequence 18, Appl

101	16	0.7	1245	4	US-09-504-358-7	Sequence 7, Appl1	174	16	0.7	10629	4	US-09-504-358-15	Sequence 15, Appl1
102	16	0.7	1245	4	US-09-954-314-7	Sequence 7, Appl1	175	16	0.7	10629	4	US-09-954-314-15	Sequence 15, Appl1
103	16	0.7	1249	3	US-08-985-950-1	Sequence 1, Appl1	176	16	0.7	11832	2	US-08-916-603-1	Sequence 1, Appl1
104	16	0.7	1314	4	US-09-134-001C-1971	Sequence 1971, Ap	177	16	0.7	12166	4	US-09-277-457-27	Sequence 27, Appl1
105	16	0.7	1332	4	US-09-134-001C-2078	Sequence 2078, Ap	178	16	0.7	13206	4	US-08-561-527-33	Sequence 33, Appl1
106	16	0.7	1467	5	PCT-US94-03437-3	Sequence 3, Appl1	179	16	0.7	13206	4	US-08-597-236-1	Sequence 1, Appl1
107	16	0.7	1500	4	US-09-134-001C-1231	Sequence 1231, Ap	180	16	0.7	14602	1	US-08-746-682A-1	Sequence 1, Appl1
108	16	0.7	1519	4	US-09-002-361-4	Sequence 4, Appl1	181	16	0.7	14602	1	US-08-561-527-113	Sequence 113, Appl1
109	16	0.7	1571	3	US-08-445-463B-3	Sequence 3, Appl1	182	16	0.7	22108	4	US-09-053-197A-7	Sequence 3, Appl1
110	16	0.7	1571	3	US-08-445-464C-3	Sequence 3, Appl1	183	16	0.7	22108	4	US-09-085-761A-3	Sequence 3, Appl1
111	16	0.7	1571	3	US-08-044-857D-3	Sequence 3, Appl1	184	16	0.7	25165	4	US-09-453-702B-39	Sequence 39, Appl1
112	16	0.7	1666	4	US-09-134-607A-8	Sequence 8, Appl1	185	16	0.7	30549	4	US-09-134-001C-132	Sequence 32, Appl1
113	16	0.7	1666	4	US-09-134-607A-13	Sequence 13, Appl1	186	16	0.7	35524	3	US-08-923-137-1	Sequence 1, Appl1
114	16	0.7	1740	4	US-09-134-607A-10	Sequence 10, Appl1	187	16	0.7	36519	3	US-08-923-137-2	Sequence 2, Appl1
115	16	0.7	1740	4	US-09-134-607A-12	Sequence 12, Appl1	188	16	0.7	37948	4	US-09-251-645-11	Sequence 11, Appl1
116	16	0.7	1740	4	US-09-134-607A-16	Sequence 16, Appl1	189	16	0.7	40328	3	US-08-742-185-102	Sequence 102, Appl1
117	16	0.7	1905	4	US-09-347-878-33	Sequence 33, Appl1	190	16	0.7	70000	4	US-09-551-896-3	Sequence 3, Appl1
118	16	0.7	1907	1	US-08-462-884A-2	Sequence 2, Appl1	191	16	0.7	98844	4	US-09-791-211-10	Sequence 10, Appl1
119	16	0.7	1908	1	US-08-461-881B-2	Sequence 2, Appl1	192	16	0.7	98844	4	US-09-791-211-10	Sequence 10, Appl1
120	16	0.7	1908	2	US-09-123-960-2	Sequence 2, Appl1	193	15	0.7	20	2	US-08-973-461A-2	Sequence 2, Appl1
121	16	0.7	2067	4	US-09-388-743-21	Sequence 21, Appl1	194	15	0.7	20	3	US-08-648-010-2	Sequence 2, Appl1
122	16	0.7	2068	4	US-09-318-448-15	Sequence 15, Appl1	195	15	0.7	22	4	US-08-542-274D-78	Sequence 78, Appl1
123	16	0.7	2266	1	US-08-453-472-1	Sequence 1, Appl1	196	15	0.7	22	4	US-08-552-014C-78	Sequence 78, Appl1
124	16	0.7	2266	1	US-08-453-952-1	Sequence 1, Appl1	197	15	0.7	27	4	US-08-584-040-6729	Sequence 6729, Ap
125	16	0.7	2266	2	US-08-484-993B-42	Sequence 42, Appl1	198	15	0.7	30	3	US-08-738-381-35	Sequence 35, Appl1
126	16	0.7	2266	2	US-08-484-158B-42	Sequence 42, Appl1	199	15	0.7	33	1	US-08-096-182A-19	Sequence 19, Appl1
127	16	0.7	2266	2	US-08-484-566A-42	Sequence 42, Appl1	200	15	0.7	33	1	US-08-096-182A-21	Sequence 21, Appl1
128	16	0.7	2266	2	US-08-484-566A-42	Sequence 42, Appl1	201	15	0.7	33	1	US-08-877-109-19	Sequence 19, Appl1
129	16	0.7	2266	3	US-08-480-150A-42	Sequence 42, Appl1	202	15	0.7	33	1	US-08-877-109-21	Sequence 21, Appl1
130	16	0.7	2266	3	US-08-458-731-42	Sequence 42, Appl1	203	15	0.7	33	3	US-08-798-760-21	Sequence 21, Appl1
131	16	0.7	2266	3	US-08-149-223A-42	Sequence 42, Appl1	204	15	0.7	33	3	PCT-US94-08327-19	Sequence 19, Appl1
132	16	0.7	2338	4	US-09-232-200-66	Sequence 66, Appl1	205	15	0.7	33	5	PCT-US94-08327-21	Sequence 21, Appl1
133	16	0.7	2338	4	US-09-232-197-66	Sequence 66, Appl1	206	15	0.7	34	2	US-08-873-479-23	Sequence 23, Appl1
134	16	0.7	2338	4	US-09-232-201-66	Sequence 66, Appl1	207	15	0.7	34	2	US-08-972-661A-23	Sequence 23, Appl1
135	16	0.7	2526	4	US-09-202-712-1	Sequence 1, Appl1	208	15	0.7	36	2	US-08-573-479-22	Sequence 22, Appl1
136	16	0.7	2539	2	US-08-248-839C-10	Sequence 10, Appl1	209	15	0.7	36	2	US-08-972-661A-22	Sequence 22, Appl1
137	16	0.7	2876	4	US-09-134-607A-9	Sequence 9, Appl1	210	15	0.7	36	4	US-09-374-962-15	Sequence 15, Appl1
138	16	0.7	2876	4	US-09-134-607A-14	Sequence 14, Appl1	211	15	0.7	36	4	US-09-374-962-16	Sequence 16, Appl1
139	16	0.7	3023	4	US-09-203-453-4	Sequence 4, Appl1	212	15	0.7	36	4	US-09-374-962-15	Sequence 15, Appl1
140	16	0.7	3066	4	US-09-066-912-1	Sequence 1, Appl1	213	15	0.7	38	3	US-08-604-757-8	Sequence 8, Appl1
141	16	0.7	3066	4	US-09-203-453-1	Sequence 1, Appl1	214	15	0.7	38	4	US-09-612-957-8	Sequence 8, Appl1
142	16	0.7	3120	4	US-09-085-761A-45	Sequence 45, Appl1	215	15	0.7	40	4	US-09-045-583-23	Sequence 23, Appl1
143	16	0.7	3236	4	US-09-318-448-23	Sequence 23, Appl1	216	15	0.7	40	4	US-09-534-185-23	Sequence 23, Appl1
144	16	0.7	3259	4	US-09-318-448-23	Sequence 23, Appl1	217	15	0.7	15	3	US-09-103-359-9	Sequence 9, Appl1
145	16	0.7	3348	4	US-09-302-620B-94	Sequence 94, Appl1	218	15	0.7	234	4	US-09-332-319-29	Sequence 29, Appl1
146	16	0.7	3510	4	US-08-961-527-53	Sequence 53, Appl1	219	15	0.7	240	1	US-08-222-177A-25	Sequence 25, Appl1
147	16	0.7	3684	2	US-08-760-075A-17	Sequence 17, Appl1	220	15	0.7	241	4	US-09-603-785-412	Sequence 412, Appl1
148	16	0.7	3684	4	US-09-338-546-17	Sequence 17, Appl1	221	15	0.7	241	4	US-09-439-313-412	Sequence 412, Appl1
149	16	0.7	3684	4	US-09-659-084-17	Sequence 17, Appl1	222	15	0.7	241	4	US-09-352-616A-412	Sequence 412, Appl1
150	16	0.7	3889	4	US-09-484-970B-39	Sequence 39, Appl1	223	15	0.7	259	3	US-09-103-359-13	Sequence 13, Appl1
151	16	0.7	4080	2	US-08-710-249-3	Sequence 3, Appl1	224	15	0.7	268	4	US-09-332-319-19	Sequence 19, Appl1
152	16	0.7	4080	4	US-09-220-157A-3	Sequence 3, Appl1	225	15	0.7	282	4	US-09-072-596-256	Sequence 256, Appl1
153	16	0.7	4117	4	US-08-961-527-100	Sequence 100, App	226	15	0.7	299	3	US-09-121-879-2	Sequence 2, Appl1
154	16	0.7	4503	2	US-08-770-301A-2	Sequence 2, Appl1	227	15	0.7	299	4	US-09-332-319-2	Sequence 2, Appl1
155	16	0.7	4503	2	US-09-175-581-2	Sequence 2, Appl1	228	15	0.7	306	4	US-08-672-213-10	Sequence 10, Appl1
156	16	0.7	4951	2	US-08-867-030B-5	Sequence 5, Appl1	229	15	0.7	306	4	US-08-672-213-45	Sequence 45, Appl1
157	16	0.7	4951	5	PCT-US95-06119-5	Sequence 5, Appl1	230	15	0.7	306	4	US-08-973-131-34	Sequence 34, Appl1
158	16	0.7	5278	4	US-08-961-527-227	Sequence 227, App	231	15	0.7	318	1	US-08-391-339-10	Sequence 10, Appl1
159	16	0.7	5537	1	US-08-135-511-32	Sequence 32, Appl1	232	15	0.7	318	1	US-07-809-437A-13	Sequence 13, Appl1
160	16	0.7	5537	1	US-08-483-852-9	Sequence 9, Appl1	233	15	0.7	318	1	US-08-476-008-10	Sequence 10, Appl1
161	16	0.7	5537	1	US-08-361-458-4	Sequence 4, Appl1	234	15	0.7	318	1	US-08-306-063-10	Sequence 10, Appl1
162	16	0.7	5537	1	US-08-477-953-9	Sequence 9, Appl1	235	15	0.7	318	1	US-08-553-943-13	Sequence 13, Appl1
163	16	0.7	5537	1	US-08-187-453-32	Sequence 32, Appl1	236	15	0.7	318	1	US-08-484-274A-10	Sequence 10, Appl1
164	16	0.7	5537	1	US-08-562-985A-5	Sequence 5, Appl1	237	15	0.7	318	1	US-08-833-485-10	Sequence 10, Appl1
165	16	0.7	5537	2	US-08-477-952-9	Sequence 9, Appl1	238	15	0.7	318	1	US-09-137-440-10	Sequence 10, Appl1
166	16	0.7	6203	4	US-09-134-218-3	Sequence 3, Appl1	239	15	0.7	318	5	PCT-US91-06146A-10	Sequence 10, Appl1
167	16	0.7	6680	4	US-09-309-572-6	Sequence 6, Appl1	240	15	0.7	318	5	PCT-US91-09437-13	Sequence 13, Appl1
168	16	0.7	8298	5	PCT-US93-03076-1	Sequence 1, Appl1	241	15	0.7	335	1	US-08-620-467A-7	Sequence 4, Appl1
169	16	0.7	8657	4	US-08-961-527-44	Sequence 44, Appl1	242	15	0.7	335	1	US-08-348-572-4	Sequence 4, Appl1
170	16	0.7	10350	2	US-09-047-125-1	Sequence 1, Appl1	243	15	0.7	335	3	US-09-041-090B-4	Sequence 4, Appl1
171	16	0.7	10350	2	US-09-047-125-2	Sequence 2, Appl1	244	15	0.7	335	4	US-08-209-603B-4	Sequence 4, Appl1
172	16	0.7	10350	3	US-07-736-335E-1	Sequence 1, Appl1	245	15	0.7	336	4	US-09-196-293-4	Sequence 4, Appl1
173	16	0.7	10350	3	US-07-736-335E-2	Sequence 2, Appl1	246	15	0.7	378	1	US-08-118-906-1	Sequence 1, Appl1



C 247	15	0.7	378	1	US-08-486-196-1	Sequence 1, Appl1	320	15	0.7	853	1	US-08-459-569-16	Sequence 16, Appl1
C 248	15	0.7	378	1	US-08-488-135-1	Sequence 1, Appl1	C 321	15	0.7	853	1	US-08-459-569-16	Sequence 16, Appl1
C 249	15	0.7	378	2	US-08-474-065-1	Sequence 1, Appl1	C 322	15	0.7	853	1	US-08-458-831-16	Sequence 16, Appl1
C 250	15	0.7	333	1	US-07-790-199-1	Sequence 1, Appl1	C 323	15	0.7	853	1	US-08-458-831-16	Sequence 16, Appl1
C 251	15	0.7	402	1	US-08-476-008-12	Sequence 12, Appl1	C 324	15	0.7	872	4	US-09-484-970B-147	Sequence 147, App
C 252	15	0.7	402	1	US-08-306-063-12	Sequence 12, Appl1	C 325	15	0.7	885	4	US-09-641-638-105	Sequence 105, App
C 253	15	0.7	402	1	US-08-833-485-12	Sequence 12, Appl1	C 326	15	0.7	895	4	US-09-602-877A-91	Sequence 91, Appl1
C 254	15	0.7	402	1	US-09-137-440-12	Sequence 12, Appl1	C 327	15	0.7	911	4	US-09-171-209-63	Sequence 63, Appl1
C 255	15	0.7	402	5	PCT-US91-06148A-12	Sequence 12, Appl1	C 328	15	0.7	924	4	US-09-171-209-63	Sequence 43, Appl1
C 256	15	0.7	427	4	US-09-397-787-12	Sequence 206, App	C 329	15	0.7	945	4	US-09-134-001C-2568	Sequence 2568, App
C 257	15	0.7	431	4	US-09-328-111-122	Sequence 122, App	C 330	15	0.7	954	3	US-09-418-641-3	Sequence 3, Appl1
C 258	15	0.7	435	4	US-09-387-787-172	Sequence 172, App	C 331	15	0.7	970	1	US-08-690-095-2	Sequence 2, Appl1
C 259	15	0.7	448	3	US-09-103-359-12	Sequence 12, Appl1	C 332	15	0.7	970	3	US-08-888-077A-28	Sequence 28, Appl1
C 260	15	0.7	450	4	US-09-397-787-173	Sequence 173, App	C 333	15	0.7	970	3	US-09-113-789-2	Sequence 2, Appl1
C 261	15	0.7	451	4	US-09-222-575-113	Sequence 113, App	C 334	15	0.7	981	4	US-09-134-001C-2097	Sequence 2097, App
C 262	15	0.7	457	4	US-08-991-789A-271	Sequence 271, App	C 335	15	0.7	993	3	US-08-669-644-21	Sequence 21, Appl1
C 263	15	0.7	457	4	US-09-062-451-271	Sequence 271, App	C 336	15	0.7	993	3	US-08-444-189-21	Sequence 21, Appl1
C 264	15	0.7	455	4	US-09-134-001C-798	Sequence 798, App	C 337	15	0.7	993	4	US-08-468-544-21	Sequence 21, Appl1
C 265	15	0.7	480	4	US-09-134-001C-1225	Sequence 1225, App	C 338	15	0.7	1002	4	US-09-641-638-591	Sequence 591, App
C 266	15	0.7	485	4	US-09-385-982-31	Sequence 31, Appl1	C 339	15	0.7	1005	3	US-09-103-359-4	Sequence 4, Appl1
C 267	15	0.7	506	4	US-09-370-838-263	Sequence 263, App	C 340	15	0.7	1039	1	US-08-458-084-9	Sequence 9, Appl1
C 268	15	0.7	519	4	US-08-580-031A-112	Sequence 12, Appl1	C 341	15	0.7	1039	1	US-08-205-508-9	Sequence 9, Appl1
C 269	15	0.7	532	4	US-09-134-001C-1189	Sequence 1189, App	C 342	15	0.7	1039	2	US-08-482-148-6	Sequence 6, Appl1
C 270	15	0.7	538	4	US-09-643-597-311	Sequence 311, App	C 343	15	0.7	1039	5	PCT-US95-02944-6	Sequence 6, Appl1
C 271	15	0.7	550	4	US-09-328-111-118	Sequence 18, Appl1	C 344	15	0.7	1039	5	PCT-US95-02945-9	Sequence 9, Appl1
C 272	15	0.7	573	4	US-08-672-213-11	Sequence 11, Appl1	C 345	15	0.7	1056	4	US-09-134-001C-2039	Sequence 2039, App
C 273	15	0.7	573	4	US-09-370-838-268	Sequence 268, App	C 346	15	0.7	1068	4	US-09-134-001C-1881	Sequence 1881, App
C 274	15	0.7	579	2	US-08-611-757-50	Sequence 50, Appl1	C 347	15	0.7	1079	2	US-08-827-190-7	Sequence 7, Appl1
C 275	15	0.7	579	5	PCT-US95-05980-50	Sequence 50, Appl1	C 348	15	0.7	1079	4	US-09-170-187-6	Sequence 7, Appl1
C 276	15	0.7	590	4	US-09-643-597-132	Sequence 132, App	C 349	15	0.7	1080	4	US-09-149-476-229	Sequence 229, App
C 277	15	0.7	591	1	US-08-686-878A-5	Sequence 5, Appl1	C 350	15	0.7	1095	4	US-08-928-383B-3	Sequence 3, Appl1
C 278	15	0.7	652	4	US-09-040-984-49	Sequence 49, Appl1	C 351	15	0.7	1097	4	US-08-858-207A-203	Sequence 203, App
C 279	15	0.7	652	4	US-09-123-912-49	Sequence 49, Appl1	C 352	15	0.7	1131	4	US-08-975-762-16	Sequence 16, Appl1
C 280	15	0.7	652	4	US-09-643-597-49	Sequence 49, Appl1	C 353	15	0.7	1131	4	US-08-821-324-16	Sequence 16, Appl1
C 281	15	0.7	676	4	US-09-605-785-355	Sequence 355, App	C 354	15	0.7	1131	4	US-09-295-028-16	Sequence 16, Appl1
C 282	15	0.7	676	4	US-09-439-113-355	Sequence 355, App	C 355	15	0.7	1131	4	US-09-106-582-16	Sequence 16, Appl1
C 283	15	0.7	676	4	US-09-352-616A-355	Sequence 355, App	C 356	15	0.7	1143	4	US-09-276-531-103	Sequence 103, App
C 284	15	0.7	679	4	US-08-642-274D-51	Sequence 51, Appl1	C 357	15	0.7	1149	4	US-09-134-001C-998	Sequence 928, App
C 285	15	0.7	679	4	US-08-952-014C-51	Sequence 51, Appl1	C 358	15	0.7	1150	2	US-08-438-439C-1	Sequence 1, Appl1
C 286	15	0.7	716	4	US-08-998-416-647	Sequence 647, App	C 359	15	0.7	1150	3	US-08-705-245-19	Sequence 19, Appl1
C 287	15	0.7	731	1	US-08-451-405A-2	Sequence 2, Appl1	C 360	15	0.7	1150	5	PCT-US96-0666B-1	Sequence 1, Appl1
C 288	15	0.7	739	4	US-09-182-117-27	Sequence 27, Appl1	C 361	15	0.7	1185	4	US-08-975-762-15	Sequence 15, Appl1
C 289	15	0.7	744	4	US-09-134-001C-2773	Sequence 2773, App	C 362	15	0.7	1185	4	US-08-821-324-15	Sequence 15, Appl1
C 290	15	0.7	750	2	US-08-602-359A-29	Sequence 29, Appl1	C 363	15	0.7	1185	4	US-09-295-028-15	Sequence 15, Appl1
C 291	15	0.7	756	4	US-09-385-982-538	Sequence 538, App	C 364	15	0.7	1185	4	US-09-106-582-15	Sequence 15, Appl1
C 292	15	0.7	768	1	US-08-464-590A-1	Sequence 1, Appl1	C 365	15	0.7	1191	4	US-09-134-001C-1743	Sequence 1743, App
C 293	15	0.7	768	3	US-09-093-585-1	Sequence 1, Appl1	C 366	15	0.7	1197	2	US-08-973-461A-3	Sequence 3, Appl1
C 294	15	0.7	773	4	US-09-020-956-3	Sequence 3, Appl1	C 367	15	0.7	1197	3	US-08-648-010-3	Sequence 3, Appl1
C 295	15	0.7	773	4	US-09-030-607-3	Sequence 3, Appl1	C 368	15	0.7	1223	3	US-08-155-005A-5	Sequence 5, Appl1
C 296	15	0.7	773	4	US-09-605-785-3	Sequence 3, Appl1	C 369	15	0.7	1223	4	US-09-363-783-5	Sequence 5, Appl1
C 297	15	0.7	773	4	US-09-439-113-3	Sequence 3, Appl1	C 370	15	0.7	1251	3	US-09-007-476-1	Sequence 1, Appl1
C 298	15	0.7	773	4	US-09-352-616A-3	Sequence 3, Appl1	C 371	15	0.7	1258	1	US-08-335-518-1	Sequence 1, Appl1
C 299	15	0.7	773	4	US-09-232-149A-3	Sequence 3, Appl1	C 372	15	0.7	1258	1	US-08-988-054-1	Sequence 1, Appl1
C 300	15	0.7	788	1	US-07-688-352C-41	Sequence 41, Appl1	C 373	15	0.7	1258	4	US-08-988-054-1	Sequence 32, Appl1
C 301	15	0.7	788	2	US-08-474-379C-41	Sequence 41, Appl1	C 374	15	0.7	1300	4	US-08-795-876-32	Sequence 31, Appl1
C 302	15	0.7	788	3	US-09-146-249A-41	Sequence 41, Appl1	C 375	15	0.7	1300	4	US-08-795-876-36	Sequence 36, Appl1
C 303	15	0.7	798	3	US-08-206-188B-41	Sequence 41, Appl1	C 376	15	0.7	1300	4	US-08-795-876-37	Sequence 37, Appl1
C 304	15	0.7	798	5	PCT-US91-02714-39	Sequence 39, Appl1	C 377	15	0.7	1300	4	US-09-134-001C-1509	Sequence 1509, App
C 305	15	0.7	801	4	US-09-134-001C-2125	Sequence 2125, App	C 378	15	0.7	1334	4	US-08-456-200B-4	Sequence 4, Appl1
C 306	15	0.7	816	4	US-09-020-956-2	Sequence 2, Appl1	C 379	15	0.7	1356	4	US-09-134-001C-617	Sequence 617, App
C 307	15	0.7	816	4	US-09-030-607-2	Sequence 2, Appl1	C 380	15	0.7	1374	4	US-09-411-687A-24	Sequence 24, Appl1
C 308	15	0.7	816	4	US-09-605-785-2	Sequence 2, Appl1	C 381	15	0.7	1377	4	US-09-221-017B-1113	Sequence 1113, App
C 309	15	0.7	816	4	US-09-439-113-2	Sequence 2, Appl1	C 382	15	0.7	1383	4	US-09-134-001C-142	Sequence 142, App
C 310	15	0.7	816	4	US-09-352-616A-2	Sequence 2, Appl1	C 383	15	0.7	1417	1	US-08-713-828-2	Sequence 2, Appl1
C 311	15	0.7	816	4	US-09-232-149A-2	Sequence 2, Appl1	C 384	15	0.7	1417	2	US-08-819-627-2	Sequence 2, Appl1
C 312	15	0.7	822	4	US-09-134-001C-804	Sequence 804, App	C 385	15	0.7	1417	2	US-09-096-245-2	Sequence 2, Appl1
C 313	15	0.7	829	3	US-08-961-083-133	Sequence 133, App	C 386	15	0.7	1429	3	US-08-155-005A-7	Sequence 7, Appl1
C 314	15	0.7	839	1	US-08-884-682-2	Sequence 2, Appl1	C 387	15	0.7	1429	4	US-09-363-783-7	Sequence 7, Appl1
C 315	15	0.7	839	2	US-09-096-082-2	Sequence 2, Appl1	C 388	15	0.7	1434	4	US-09-480-921B-9	Sequence 9, Appl1
C 316	15	0.7	841	2	US-08-998-416-318	Sequence 318, App	C 389	15	0.7	1448	4	US-08-680-921B-28	Sequence 28, App
C 317	15	0.7	850	2	US-08-560-398-7	Sequence 7, Appl1	C 390	15	0.7	1448	1	US-08-585-758A-1	Sequence 1, Appl1
C 318	15	0.7	853	1	US-08-008-216-16	Sequence 16, Appl1	C 391	15	0.7	1448	1	US-08-977-818-1	Sequence 1, Appl1
C 319	15	0.7	853	1	US-08-008-216-16	Sequence 16, Appl1	C 392	15	0.7	1448	2	US-08-670-274B-1	Sequence 1, Appl1

393	15	0.7	1448	4	US-09-146-187-1	Sequence 1, Appl1	C 466	15	0.7	1712	4	US-08-809-326A-8	Sequence 8, Appl1
394	15	0.7	1450	2	US-08-692-787-6	Sequence 6, Appl1	C 467	15	0.7	1729	3	US-08-155-005A-16	Sequence 16, Appl1
395	15	0.7	1452	1	US-08-149-093A-3	Sequence 3, Appl1	C 468	15	0.7	1729	4	US-09-363-783-16	Sequence 16, Appl1
396	15	0.7	1452	1	US-08-911-245-3	Sequence 3, Appl1	C 469	15	0.7	1722	4	US-09-449-335-1	Sequence 1, Appl1
397	15	0.7	1452	1	US-08-553-058C-3	Sequence 3, Appl1	C 470	15	0.7	1732	4	US-09-449-335-5	Sequence 5, Appl1
398	15	0.7	1452	2	US-08-514-451A-3	Sequence 3, Appl1	C 471	15	0.7	1744	3	US-08-961-083-131	Sequence 131, Appl1
399	15	0.7	1452	3	US-09-170-311-3	Sequence 3, Appl1	C 472	15	0.7	1758	4	US-09-370-838-25	Sequence 25, Appl1
400	15	0.7	1452	3	US-09-510-473-3	Sequence 3, Appl1	C 473	15	0.7	1762	1	US-08-105-483-300	Sequence 300, Appl1
401	15	0.7	1452	4	US-09-048-916B-3	Sequence 3, Appl1	C 474	15	0.7	1762	1	US-08-709-209-300	Sequence 300, Appl1
402	15	0.7	1456	4	US-09-056-285A-6	Sequence 6, Appl1	C 475	15	0.7	1782	1	US-08-458-101-300	Sequence 300, Appl1
403	15	0.7	1458	4	US-09-480-921B-3	Sequence 3, Appl1	C 476	15	0.7	1783	4	US-09-312-285-4	Sequence 4, Appl1
404	15	0.7	1458	4	US-09-417-359A-4	Sequence 4, Appl1	C 477	15	0.7	1783	4	US-09-312-266-4	Sequence 2, Appl1
405	15	0.7	1468	4	US-09-290-734-23	Sequence 23, Appl1	C 478	15	0.7	1763	4	US-09-312-038-2	Sequence 4, Appl1
406	15	0.7	1468	4	US-09-290-734-25	Sequence 25, Appl1	C 479	15	0.7	1763	4	US-09-728-764-4	Sequence 4, Appl1
407	15	0.7	1473	4	US-09-056-285A-9	Sequence 9, Appl1	C 480	15	0.7	1753	4	US-09-312-304B-3	Sequence 3, Appl1
408	15	0.7	1484	4	US-09-210-168-2	Sequence 2, Appl1	C 481	15	0.7	1763	4	US-09-728-792-4	Sequence 4, Appl1
409	15	0.7	1484	4	US-09-134-001C-1436	Sequence 1436, Appl1	C 482	15	0.7	1767	3	US-09-199-859-1	Sequence 1, Appl1
410	15	0.7	1494	4	US-09-413-611A-5	Sequence 5, Appl1	C 483	15	0.7	1782	1	US-08-153-799-15	Sequence 15, Appl1
411	15	0.7	1500	4	US-09-413-611A-7	Sequence 7, Appl1	C 484	15	0.7	1782	4	US-08-769-746-1	Sequence 1, Appl1
412	15	0.7	1512	4	US-09-134-001C-643	Sequence 643, Appl1	C 485	15	0.7	1792	3	US-08-747-221B-18	Sequence 18, Appl1
413	15	0.7	1515	3	US-08-747-221B-16	Sequence 16, Appl1	C 486	15	0.7	1792	3	US-08-747-221B-20	Sequence 20, Appl1
414	15	0.7	1515	3	US-08-747-221B-17	Sequence 17, Appl1	C 487	15	0.7	1782	4	US-09-005-051-18	Sequence 18, Appl1
415	15	0.7	1515	3	US-09-005-051-16	Sequence 16, Appl1	C 488	15	0.7	1782	4	US-09-005-051-20	Sequence 20, Appl1
416	15	0.7	1515	4	US-09-005-051-17	Sequence 17, Appl1	C 489	15	0.7	1792	4	US-09-491-522-4	Sequence 4, Appl1
417	15	0.7	1515	4	US-08-032-869A-1	Sequence 1, Appl1	C 490	15	0.7	1807	1	US-08-118-906-13	Sequence 13, Appl1
418	15	0.7	1517	1	US-08-472-673-1	Sequence 2, Appl1	C 491	15	0.7	1807	1	US-08-488-135-13	Sequence 13, Appl1
419	15	0.7	1517	1	US-08-472-673-2	Sequence 2, Appl1	C 492	15	0.7	1807	1	US-08-488-135-13	Sequence 13, Appl1
420	15	0.7	1517	1	US-08-472-673-2	Sequence 2, Appl1	C 493	15	0.7	1807	2	US-08-474-065-13	Sequence 13, Appl1
421	15	0.7	1517	1	US-08-475-782-1	Sequence 1, Appl1	C 494	15	0.7	1825	4	US-09-184-964-1	Sequence 1, Appl1
422	15	0.7	1517	1	US-08-475-782-2	Sequence 2, Appl1	C 495	15	0.7	1827	4	US-08-897-956A-5	Sequence 5, Appl1
423	15	0.7	1517	2	US-08-472-678-1	Sequence 2, Appl1	C 496	15	0.7	1830	1	US-07-691-079C-3	Sequence 3, Appl1
424	15	0.7	1517	2	US-08-472-678-2	Sequence 2, Appl1	C 497	15	0.7	1830	1	US-08-433-037-3	Sequence 3, Appl1
425	15	0.7	1517	2	US-08-484-503-1	Sequence 1, Appl1	C 498	15	0.7	1830	2	US-08-702-572-1	Sequence 1, Appl1
426	15	0.7	1517	2	US-08-484-503-2	Sequence 2, Appl1	C 499	15	0.7	1836	1	US-07-929-580B-1	Sequence 1, Appl1
427	15	0.7	1518	4	US-08-944-368A-1	Sequence 1, Appl1	C 500	15	0.7	1839	1	US-07-929-580B-4	Sequence 4, Appl1
428	15	0.7	1518	4	US-09-820-764-1	Sequence 1, Appl1	C 501	15	0.7	1850	3	US-08-724-466B-3	Sequence 3, Appl1
429	15	0.7	1528	4	US-09-362-473-13	Sequence 13, Appl1	C 502	15	0.7	1850	4	US-08-882-164B-3	Sequence 3, Appl1
430	15	0.7	1537	3	US-09-318-448-32	Sequence 32, Appl1	C 503	15	0.7	1855	4	US-08-810-009-1	Sequence 1, Appl1
431	15	0.7	1567	3	US-08-889-108-16	Sequence 16, Appl1	C 504	15	0.7	1862	2	US-08-797-689-1	Sequence 1, Appl1
432	15	0.7	1567	5	PCT-US94-10358-16	Sequence 16, Appl1	C 505	15	0.7	1863	2	US-08-455-072B-3	Sequence 3, Appl1
433	15	0.7	1584	4	US-08-928-383B-1	Sequence 1, Appl1	C 506	15	0.7	1863	2	US-09-311-626B-5	Sequence 5, Appl1
434	15	0.7	1590	3	US-08-747-221B-23	Sequence 23, Appl1	C 507	15	0.7	1874	4	US-09-167-299-1	Sequence 1, Appl1
435	15	0.7	1590	4	US-09-005-051-23	Sequence 23, Appl1	C 508	15	0.7	1890	5	US-09-360-545-77	Sequence 77, Appl1
436	15	0.7	1598	4	US-09-661-711A-7	Sequence 7, Appl1	C 509	15	0.7	1894	5	PCT-US91-08177-10	Sequence 10, Appl1
437	15	0.7	1615	3	US-09-040-508-1	Sequence 1, Appl1	C 510	15	0.7	1902	2	US-07-688-352C-43	Sequence 43, Appl1
438	15	0.7	1615	4	US-09-500-654-1	Sequence 1, Appl1	C 511	15	0.7	1902	2	US-08-474-379C-43	Sequence 43, Appl1
439	15	0.7	1621	4	US-09-020-956-107	Sequence 107, Appl1	C 512	15	0.7	1902	3	US-09-146-249A-43	Sequence 43, Appl1
440	15	0.7	1621	4	US-09-030-607-107	Sequence 107, Appl1	C 513	15	0.7	1902	3	US-08-206-188B-43	Sequence 43, Appl1
441	15	0.7	1621	4	US-09-605-785-107	Sequence 107, Appl1	C 514	15	0.7	1902	5	PCT-US91-02714-40	Sequence 40, Appl1
442	15	0.7	1621	4	US-09-439-313-107	Sequence 107, Appl1	C 515	15	0.7	1939	1	US-08-212-188-3	Sequence 3, Appl1
443	15	0.7	1621	4	US-09-352-616A-107	Sequence 107, Appl1	C 516	15	0.7	1939	3	US-08-970-725-3	Sequence 3, Appl1
444	15	0.7	1621	2	US-09-232-149A-107	Sequence 107, Appl1	C 517	15	0.7	1939	5	PCT-US95-02708-3	Sequence 3, Appl1
445	15	0.7	1630	3	US-08-805-191-1	Sequence 1, Appl1	C 518	15	0.7	1944	4	US-09-377-557-19	Sequence 19, Appl1
446	15	0.7	1630	3	US-08-747-221B-21	Sequence 21, Appl1	C 519	15	0.7	1944	4	US-08-809-326A-13	Sequence 13, Appl1
447	15	0.7	1650	3	US-08-747-221B-22	Sequence 22, Appl1	C 520	15	0.7	1962	3	US-08-747-221B-13	Sequence 13, Appl1
448	15	0.7	1650	3	US-08-747-221B-60	Sequence 60, Appl1	C 521	15	0.7	1982	3	US-08-747-221B-15	Sequence 15, Appl1
449	15	0.7	1650	3	US-08-747-221B-61	Sequence 61, Appl1	C 522	15	0.7	1982	4	US-09-005-051-13	Sequence 13, Appl1
450	15	0.7	1650	4	US-09-005-051-21	Sequence 21, Appl1	C 523	15	0.7	1982	4	US-09-005-051-15	Sequence 15, Appl1
451	15	0.7	1650	4	US-09-005-051-22	Sequence 22, Appl1	C 524	15	0.7	1987	2	US-08-990-379-1	Sequence 1, Appl1
452	15	0.7	1650	4	US-09-005-051-60	Sequence 60, Appl1	C 525	15	0.7	1991	4	US-08-235-836C-69	Sequence 69, Appl1
453	15	0.7	1650	4	US-09-005-051-61	Sequence 61, Appl1	C 526	15	0.7	1991	4	US-08-235-836C-77	Sequence 77, Appl1
454	15	0.7	1651	3	US-08-155-005A-3	Sequence 3, Appl1	C 527	15	0.7	1993	2	US-08-990-379-2	Sequence 2, Appl1
455	15	0.7	1651	4	US-09-363-783-3	Sequence 3, Appl1	C 528	15	0.7	2000	1	US-09-041-075A-4	Sequence 4, Appl1
456	15	0.7	1664	4	US-09-045-583-4	Sequence 4, Appl1	C 529	15	0.7	2000	1	US-09-041-075A-6	Sequence 6, Appl1
457	15	0.7	1664	4	US-09-534-185-4	Sequence 4, Appl1	C 530	15	0.7	2032	4	US-09-500-063-1	Sequence 1, Appl1
458	15	0.7	1670	4	US-09-071-035-267	Sequence 267, Appl1	C 531	15	0.7	2032	4	US-08-604-789B-12	Sequence 12, Appl1
459	15	0.7	1670	5	PCT-US94-01321-71	Sequence 71, Appl1	C 532	15	0.7	2040	3	US-08-604-789B-12	Sequence 12, Appl1
460	15	0.7	1677	1	US-08-285-440-12	Sequence 12, Appl1	C 533	15	0.7	2040	4	US-09-312-721B-11	Sequence 11, Appl1
461	15	0.7	1677	1	US-08-610-349-12	Sequence 12, Appl1	C 534	15	0.7	2040	4	US-09-312-721B-12	Sequence 12, Appl1
462	15	0.7	1693	1	US-09-661-711A-1	Sequence 1, Appl1	C 535	15	0.7	2064	1	US-08-164-839-30	Sequence 30, Appl1
463	15	0.7	1697	4	US-09-556-870A-1	Sequence 1, Appl1	C 536	15	0.7	2064	1	US-08-164-839-32	Sequence 32, Appl1
464	15	0.7	1699	4	US-09-661-711A-9	Sequence 9, Appl1	C 537	15	0.7	2064	1	US-08-583-799-30	Sequence 30, Appl1
465	15	0.7	1702	4	US-08-934-254-26	Sequence 26, Appl1	C 538	15	0.7	2064	1	US-08-583-799-32	Sequence 32, Appl1

C 539	15	0.7	2104	1	US-08-592-126-96	Sequence 96, Appl	612	15	0.7	2499	4	US-09-276-531-133	Sequence 133, App
C 540	15	0.7	2107	4	US-08-235-836C-73	Sequence 72, Appl	613	15	0.7	2505	2	US-08-890-094-1	Sequence 1, Appl1
C 541	15	0.7	2112	4	US-09-232-160-12	Sequence 12, Appl	614	15	0.7	2547	3	US-08-508-761B-1	Sequence 1, Appl1
C 542	15	0.7	2129	4	US-08-975-762-39	Sequence 39, Appl	615	15	0.7	2577	4	US-09-605-785-552	Sequence 552, App
C 543	15	0.7	2129	4	US-09-295-028-39	Sequence 39, Appl	616	15	0.7	2601	3	US-09-039-773A-2	Sequence 2, Appl1
C 544	15	0.7	2129	4	US-09-106-582-39	Sequence 39, Appl	617	15	0.7	2606	4	US-09-252-816A-7	Sequence 7, Appl1
C 545	15	0.7	2144	3	US-08-747-221B-57	Sequence 57, Appl	618	15	0.7	2628	2	US-08-696-944-1	Sequence 1, Appl1
C 546	15	0.7	2144	3	US-08-747-221B-59	Sequence 59, Appl	619	15	0.7	2640	4	US-09-312-285-3	Sequence 3, Appl1
C 547	15	0.7	2144	4	US-09-005-051-57	Sequence 57, Appl	620	15	0.7	2640	4	US-09-312-266-3	Sequence 3, Appl1
C 548	15	0.7	2144	4	US-09-005-051-59	Sequence 59, Appl	621	15	0.7	2640	4	US-09-312-038-1	Sequence 1, Appl1
C 549	15	0.7	2148	1	US-08-164-839-69	Sequence 69, Appl	622	15	0.7	2640	4	US-09-128-764-3	Sequence 3, Appl1
C 550	15	0.7	2148	1	US-08-164-839-71	Sequence 71, Appl	623	15	0.7	2640	4	US-09-312-304B-1	Sequence 1, Appl1
C 551	15	0.7	2148	1	US-08-583-799-69	Sequence 69, Appl	624	15	0.7	2706	2	US-09-728-792-3	Sequence 3, Appl1
C 552	15	0.7	2148	1	US-08-583-799-71	Sequence 71, Appl	625	15	0.7	2706	2	US-08-454-549-1	Sequence 1, Appl1
C 553	15	0.7	2170	4	US-09-480-921B-14	Sequence 14, Appl	626	15	0.7	2706	3	US-08-454-552-1	Sequence 1, Appl1
C 554	15	0.7	2178	2	US-08-474-379C-80	Sequence 80, Appl	627	15	0.7	2706	3	US-08-454-552-1	Sequence 1, Appl1
C 555	15	0.7	2178	3	US-09-146-249A-80	Sequence 80, Appl	628	15	0.7	2712	1	US-08-346-455B-37	Sequence 37, Appl
C 556	15	0.7	2178	3	US-08-206-188B-80	Sequence 80, Appl	629	15	0.7	2712	3	US-08-977-221-37	Sequence 37, Appl
C 557	15	0.7	2179	1	US-08-487-886-1	Sequence 1, Appl1	630	15	0.7	2712	4	US-09-483-831B-37	Sequence 37, Appl
C 558	15	0.7	2179	2	US-08-531-070A-1	Sequence 1, Appl1	631	15	0.7	2712	5	PCT-US95-06613-37	Sequence 37, Appl
C 559	15	0.7	2179	3	US-08-482-855-1	Sequence 1, Appl1	632	15	0.7	2728	4	US-09-232-160-1	Sequence 1, Appl1
C 560	15	0.7	2179	4	US-08-474-886-1	Sequence 1, Appl1	633	15	0.7	2728	4	US-09-643-597-358	Sequence 1, Appl1
C 561	15	0.7	2181	4	US-09-239-041-1	Sequence 1, Appl1	634	15	0.7	2773	4	US-09-643-597-358	Sequence 358, App
C 562	15	0.7	2192	1	US-08-339-696-125	Sequence 125, App	635	15	0.7	2784	4	US-09-643-597-358	Sequence 168, App
C 563	15	0.7	2193	4	US-09-427-261-2	Sequence 2, Appl1	636	15	0.7	2875	3	US-09-455-893-1	Sequence 1, Appl1
C 564	15	0.7	2193	4	US-09-427-261-3	Sequence 3, Appl1	637	15	0.7	2935	4	US-08-897-956A-4	Sequence 4, Appl1
C 565	15	0.7	2196	1	US-08-090-523-29	Sequence 29, Appl	638	15	0.7	2970	4	US-09-193-562D-31	Sequence 31, Appl
C 566	15	0.7	2196	1	US-08-334-639-9	Sequence 1, Appl1	639	15	0.7	2995	1	US-08-592-126-85	Sequence 85, Appl
C 567	15	0.7	2196	1	US-08-338-627-29	Sequence 29, Appl	640	15	0.7	3095	6	US-08-680-326-27	Sequence 27, Appl
C 568	15	0.7	2196	1	US-08-406-858-94	Sequence 29, Appl	641	15	0.7	3095	6	5231168-1	Sequence 27, Appl
C 569	15	0.7	2196	5	PCT-US94-05275-24	Sequence 24, Appl	642	15	0.7	3104	1	US-08-346-455B-66	Sequence 66, Appl
C 570	15	0.7	2214	2	US-08-484-993B-1	Sequence 1, Appl1	643	15	0.7	3104	3	US-09-483-831B-66	Sequence 66, Appl
C 571	15	0.7	2214	2	US-08-484-158B-1	Sequence 1, Appl1	644	15	0.7	3104	4	US-09-483-831B-66	Sequence 66, Appl
C 572	15	0.7	2214	2	US-08-484-596A-1	Sequence 1, Appl1	645	15	0.7	3104	5	PCT-US95-06613-66	Sequence 66, Appl
C 573	15	0.7	2214	2	US-08-480-150A-1	Sequence 1, Appl1	646	15	0.7	3134	4	US-09-134-001C-2528	Sequence 2528, App
C 574	15	0.7	2214	3	US-08-458-731-1	Sequence 1, Appl1	647	15	0.7	3130	2	US-08-744-379C-62	Sequence 62, Appl
C 575	15	0.7	2214	3	US-08-149-223A-1	Sequence 1, Appl1	648	15	0.7	3130	3	US-09-146-249A-62	Sequence 62, Appl
C 576	15	0.7	2214	4	US-08-943-731-57	Sequence 57, Appl	649	15	0.7	3130	3	US-08-206-188B-62	Sequence 62, Appl
C 577	15	0.7	2215	1	US-08-285-440-14	Sequence 14, Appl	650	15	0.7	3140	4	US-08-661-527-158	Sequence 158, App
C 578	15	0.7	2215	1	US-08-630-349-14	Sequence 14, Appl	651	15	0.7	3143	1	US-08-485-621-1	Sequence 1, Appl1
C 579	15	0.7	2235	4	US-09-484-370B-28	Sequence 28, Appl	652	15	0.7	3143	5	US-08-973-831-1	Sequence 1, Appl1
C 580	15	0.7	2235	4	US-08-46-508-1	Sequence 1, Appl1	653	15	0.7	3143	5	PCT-US96-09530A-1	Sequence 1, Appl1
C 581	15	0.7	2248	1	US-08-369-780-1	Sequence 1, Appl1	654	15	0.7	3153	4	US-09-175-928-9	Sequence 9, Appl1
C 582	15	0.7	2248	1	US-08-475-682-1	Sequence 1, Appl1	655	15	0.7	3182	4	US-09-221-017B-1029	Sequence 1029, App
C 583	15	0.7	2248	1	US-08-780-833-1	Sequence 1, Appl1	656	15	0.7	3251	3	US-08-346-455B-68	Sequence 68, Appl
C 584	15	0.7	2248	1	US-08-636-036-1	Sequence 1, Appl1	657	15	0.7	3251	3	US-08-977-221-68	Sequence 68, Appl
C 585	15	0.7	2248	3	US-08-918-509-1	Sequence 1, Appl1	658	15	0.7	3251	4	US-09-483-831B-68	Sequence 68, Appl
C 586	15	0.7	2248	3	US-09-108-362-1	Sequence 1, Appl1	659	15	0.7	3251	5	PCT-US95-06613-68	Sequence 68, Appl
C 587	15	0.7	2248	3	US-08-899-336-2	Sequence 2, Appl1	660	15	0.7	3353	1	US-08-453-695A-109	Sequence 109, App
C 588	15	0.7	2248	3	US-08-899-336-2	Sequence 2, Appl1	661	15	0.7	3353	2	US-08-453-695A-109	Sequence 109, App
C 589	15	0.7	2248	3	US-08-899-336-2	Sequence 2, Appl1	662	15	0.7	3353	2	US-08-453-695A-109	Sequence 109, App
C 590	15	0.7	2248	3	US-08-899-336-2	Sequence 2, Appl1	663	15	0.7	3353	4	US-09-099-639-109	Sequence 109, App
C 591	15	0.7	2248	3	US-08-250-740-32	Sequence 32, Appl	664	15	0.7	3353	5	PCT-US95-08071-109	Sequence 109, App
C 592	15	0.7	2248	3	US-07-695-472B-34	Sequence 34, Appl	665	15	0.7	3353	5	US-09-643-597-167	Sequence 167, App
C 593	15	0.7	2248	3	US-07-695-472B-34	Sequence 34, Appl	666	15	0.7	3353	4	US-09-146-969-4	Sequence 4, Appl1
C 594	15	0.7	2248	3	US-07-695-472B-34	Sequence 34, Appl	667	15	0.7	3353	4	US-09-252-816A-6	Sequence 6, Appl1
C 595	15	0.7	2248	3	US-07-695-472B-34	Sequence 34, Appl	668	15	0.7	3353	4	US-09-026-033-17	Sequence 17, Appl
C 596	15	0.7	2248	3	US-07-695-472B-34	Sequence 34, Appl	669	15	0.7	3353	4	US-09-221-017B-106	Sequence 106, App
C 597	15	0.7	2248	3	US-07-695-472B-34	Sequence 34, Appl	670	15	0.7	3353	3	US-08-714-918-7	Sequence 7, Appl1
C 598	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	671	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 599	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	672	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 600	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	673	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 601	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	674	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 602	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	675	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 603	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	676	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 604	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	677	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 605	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	678	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 606	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	679	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 607	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	680	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 608	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	681	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 609	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	682	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 610	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	683	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1
C 611	15	0.7	2248	3	US-08-484-993B-9	Sequence 9, Appl1	684	15	0.7	3353	4	US-09-265-315-7	Sequence 7, Appl1

C 685	15	0.7	3587	2	US-08-874-186-91	Sequence 91, Appl	758	15	0.7	5433	3	US-09-156-842-35	Sequence 35, Appl
C 686	15	0.7	3596	2	US-08-779-801-5	Sequence 5, Appl	759	15	0.7	5458	4	US-09-199-865-2	Sequence 2, Appl
C 687	15	0.7	3596	4	US-09-298-441-5	Sequence 5, Appl	760	15	0.7	5461	3	US-07-998-2898-7	Sequence 7, Appl
C 688	15	0.7	3632	2	US-08-779-801-3	Sequence 3, Appl	761	15	0.7	5474	4	US-09-400-738-1	Sequence 1, Appl
C 689	15	0.7	3632	2	US-08-779-801-4	Sequence 3, Appl	762	15	0.7	5474	4	US-08-652-4264-1	Sequence 1, Appl
C 690	15	0.7	3632	4	US-09-298-441-3	Sequence 4, Appl	763	15	0.7	5499	1	US-07-695-564-4	Sequence 4, Appl
C 691	15	0.7	3632	4	US-09-298-441-4	Sequence 4, Appl	764	15	0.7	5499	1	US-08-241-387-4	Sequence 4, Appl
C 692	15	0.7	3796	1	US-08-920-812-19	Sequence 19, Appl	765	15	0.7	5500	3	US-08-882-046-1	Sequence 1, Appl
C 693	15	0.7	3796	1	US-08-920-827-19	Sequence 19, Appl	766	15	0.7	5529	1	US-07-695-564-2	Sequence 2, Appl
C 694	15	0.7	3796	1	US-08-921-177-19	Sequence 19, Appl	767	15	0.7	5629	1	US-08-241-387-1	Sequence 2, Appl
C 695	15	0.7	3796	1	US-08-362-577C-19	Sequence 19, Appl	768	15	0.7	5668	4	US-09-293-170-7	Sequence 7, Appl
C 696	15	0.7	3796	2	US-08-920-828-19	Sequence 19, Appl	769	15	0.7	5761	4	US-09-323-4724-1	Sequence 1, Appl
C 697	15	0.7	3804	3	US-09-356-952-14	Sequence 14, Appl	770	15	0.7	5761	4	US-09-323-4724-3	Sequence 3, Appl
C 698	15	0.7	3871	4	US-09-484-970B-70	Sequence 70, Appl	771	15	0.7	5762	4	US-09-323-4724-11	Sequence 11, Appl
C 699	15	0.7	3935	4	US-09-060-482-1	Sequence 1, Appl	772	15	0.7	5762	4	US-09-323-4724-13	Sequence 13, Appl
C 700	15	0.7	3951	4	US-09-643-597-160	Sequence 160, App	773	15	0.7	5816	4	US-08-857-076-11	Sequence 11, Appl
C 701	15	0.7	3991	4	US-08-506-296B-3	Sequence 3, Appl	774	15	0.7	5837	1	US-07-686-340-1	Sequence 1, Appl
C 702	15	0.7	3998	4	US-09-066-046-5	Sequence 5, Appl	775	15	0.7	5837	1	US-08-004-139B-1	Sequence 1, Appl
C 703	15	0.7	4004	4	US-09-293-505-8	Sequence 8, Appl	776	15	0.7	5837	1	US-08-117-491-1	Sequence 1, Appl
C 704	15	0.7	4030	4	US-09-293-505-1	Sequence 1, Appl	777	15	0.7	5837	1	US-08-271-364A-1	Sequence 1, Appl
C 705	15	0.7	4160	4	US-09-341-678-5	Sequence 5, Appl	778	15	0.7	5837	2	US-08-811-492-1	Sequence 1, Appl
C 706	15	0.7	4162	2	US-08-459-448A-26	Sequence 26, Appl	779	15	0.7	5837	2	US-08-322-715B-1	Sequence 1, Appl
C 707	15	0.7	4162	3	US-08-459-595A-26	Sequence 26, Appl	780	15	0.7	5837	5	PCT-US96-10545A-1	Sequence 1, Appl
C 708	15	0.7	4162	3	US-08-459-504B-26	Sequence 26, Appl	C 781	15	0.7	5856	4	US-09-293-170-1	Sequence 1, Appl
C 709	15	0.7	4162	3	US-08-459-444-26	Sequence 26, Appl	782	15	0.7	5856	3	US-08-665-259-24	Sequence 1, Appl
C 710	15	0.7	4162	4	US-09-547-422-26	Sequence 26, Appl	783	15	0.7	5894	3	US-08-762-500-24	Sequence 24, Appl
C 711	15	0.7	4164	1	US-08-188-582-3	Sequence 3, Appl	784	15	0.7	5912	2	US-08-629-001A-1	Sequence 1, Appl
C 712	15	0.7	4164	1	US-08-646-715-3	Sequence 3, Appl	785	15	0.7	5912	4	US-08-642-2740-1	Sequence 1, Appl
C 713	15	0.7	4165	1	US-07-951-715A-26	Sequence 26, Appl	786	15	0.7	5912	4	US-08-952-127-1	Sequence 1, Appl
C 714	15	0.7	4174	4	US-09-221-017B-926	Sequence 926, App	787	15	0.7	5912	4	US-08-952-014C-1	Sequence 1, Appl
C 715	15	0.7	4208	4	US-09-214-278-6	Sequence 6, Appl	788	15	0.7	6226	1	US-08-542-363-1	Sequence 1, Appl
C 716	15	0.7	4308	4	US-09-068-740A-10	Sequence 10, Appl	789	15	0.7	6226	4	US-09-100-089-1	Sequence 1, Appl
C 717	15	0.7	4391	4	US-09-207-857-1	Sequence 1, Appl	790	15	0.7	6226	4	US-09-670-827-1	Sequence 1, Appl
C 718	15	0.7	4434	3	US-08-815-809-1	Sequence 1, Appl	C 791	15	0.7	6273	4	US-08-961-527-21	Sequence 21, Appl
C 719	15	0.7	4440	1	US-08-200-016-4	Sequence 4, Appl	792	15	0.7	6454	2	US-08-400-159-5	Sequence 5, Appl
C 720	15	0.7	4440	3	US-07-792-600-1	Sequence 1, Appl	793	15	0.7	6454	3	US-08-611-729A-5	Sequence 5, Appl
C 721	15	0.7	4440	3	US-09-157-021-1	Sequence 1, Appl	794	15	0.7	6513	1	US-08-338-702-7	Sequence 7, Appl
C 722	15	0.7	4440	3	US-09-156-842-1	Sequence 1, Appl	795	15	0.7	6513	1	US-08-337-339-7	Sequence 7, Appl
C 723	15	0.7	4467	1	US-08-565-907A-1	Sequence 1, Appl	796	15	0.7	6513	1	US-08-724-095-7	Sequence 7, Appl
C 724	15	0.7	4467	1	US-08-565-907B-1	Sequence 1, Appl	797	15	0.7	6513	5	PCT-US95-14267-7	Sequence 7, Appl
C 725	15	0.7	4467	2	US-08-910-551B-1	Sequence 1, Appl	798	15	0.7	6513	5	PCT-US95-14378-7	Sequence 7, Appl
C 726	15	0.7	4467	2	US-08-910-551B-1	Sequence 1, Appl	799	15	0.7	6519	2	US-08-808-793-24	Sequence 24, Appl
C 727	15	0.7	4467	2	US-08-909-425A-1	Sequence 1, Appl	800	15	0.7	6525	1	US-08-493-092-3	Sequence 3, Appl
C 728	15	0.7	4467	2	US-08-909-425A-1	Sequence 1, Appl	801	15	0.7	6525	1	US-08-508-836A-3	Sequence 3, Appl
C 729	15	0.7	4472	2	US-08-816-155B-3	Sequence 3, Appl	802	15	0.7	6525	3	US-08-762-500-74	Sequence 74, Appl
C 730	15	0.7	4472	2	US-09-079-587-3	Sequence 3, Appl	803	15	0.7	6587	4	US-09-293-170-8	Sequence 8, Appl
C 731	15	0.7	4517	5	PCT-US93-06251-83	Sequence 83, Appl	804	15	0.7	6632	4	US-09-491-522-1	Sequence 1, Appl
C 732	15	0.7	4539	1	US-08-119-512-1	Sequence 1, Appl	805	15	0.7	6709	4	US-09-285-601-3	Sequence 3, Appl
C 733	15	0.7	4539	1	US-08-488-015B-1	Sequence 1, Appl	806	15	0.7	6827	4	US-08-961-527-60	Sequence 60, Appl
C 734	15	0.7	4545	3	US-08-814-412-11	Sequence 11, Appl	807	15	0.7	6924	1	US-08-015-973-2	Sequence 2, Appl
C 735	15	0.7	4545	3	US-09-221-017B-1011	Sequence 1011, Ap	808	15	0.7	6924	2	US-08-448-164-2	Sequence 2, Appl
C 736	15	0.7	4615	1	US-08-188-582-1	Sequence 1, Appl	809	15	0.7	6942	2	US-08-460-309-3	Sequence 3, Appl
C 737	15	0.7	4615	1	US-08-646-715-1	Sequence 1, Appl	810	15	0.7	6942	2	US-08-125-077-3	Sequence 3, Appl
C 738	15	0.7	4624	4	US-09-066-046-3	Sequence 3, Appl	811	15	0.7	7001	4	US-08-961-527-88	Sequence 88, Appl
C 739	15	0.7	4741	1	US-07-695-472B-4	Sequence 4, Appl	812	15	0.7	7233	5	PCT-US91-08177-1	Sequence 1, Appl
C 740	15	0.7	4742	1	US-08-250-740-35	Sequence 35, Appl	813	15	0.7	7330	1	US-07-865-662B-14	Sequence 14, Appl
C 741	15	0.7	4757	4	US-07-865-662F-12	Sequence 12, Appl	814	15	0.7	7330	4	US-08-374-219B-16	Sequence 16, Appl
C 742	15	0.7	4757	4	US-08-374-219B-12	Sequence 12, Appl	815	15	0.7	7408	1	US-08-441-822-1	Sequence 1, Appl
C 743	15	0.7	4792	4	US-08-961-527-156	Sequence 156, App	816	15	0.7	7410	1	US-08-493-092-1	Sequence 1, Appl
C 744	15	0.7	4977	4	US-09-071-035-257	Sequence 257, App	817	15	0.7	7410	1	US-08-508-836A-1	Sequence 1, Appl
C 745	15	0.7	4977	4	US-09-071-035-261	Sequence 261, App	818	15	0.7	7502	3	US-08-969-644-6	Sequence 6, Appl
C 746	15	0.7	4977	4	US-09-071-035-265	Sequence 265, App	819	15	0.7	7502	3	US-08-444-189-6	Sequence 6, Appl
C 747	15	0.7	5014	4	US-09-381-862-3	Sequence 3, Appl	820	15	0.7	7502	4	US-08-668-544-6	Sequence 6, Appl
C 748	15	0.7	5057	3	US-08-651-989A-6	Sequence 6, Appl	821	15	0.7	7521	4	US-09-004-838-116	Sequence 116, App
C 749	15	0.7	5057	4	US-09-385-752-6	Sequence 6, Appl	822	15	0.7	7577	4	US-08-961-527-46	Sequence 46, Appl
C 750	15	0.7	5063	4	US-08-185-432-1	Sequence 1, Appl	823	15	0.7	7685	4	US-09-221-017B-1092	Sequence 1092, Ap
C 751	15	0.7	5174	4	US-09-066-047-9	Sequence 9, Appl	824	15	0.7	7791	2	US-08-149-0970-23	Sequence 23, Appl
C 752	15	0.7	5186	2	US-08-596-366-1	Sequence 1, Appl	825	15	0.7	7791	3	US-08-949-386-23	Sequence 23, Appl
C 753	15	0.7	5186	2	US-08-967-104-1	Sequence 1, Appl	826	15	0.7	7791	3	US-08-450-562-23	Sequence 23, Appl
C 754	15	0.7	5224	2	US-08-874-186-46	Sequence 46, Appl	827	15	0.7	7791	4	US-08-984-709B-23	Sequence 23, Appl
C 755	15	0.7	5375	2	US-08-757-223-7	Sequence 7, Appl	828	15	0.7	7791	4	US-08-450-272-23	Sequence 23, Appl
C 756	15	0.7	5402	4	US-09-221-017B-194	Sequence 194, App	829	15	0.7	7808	2	US-08-149-0970-22	Sequence 22, Appl
C 757	15	0.7	5433	3	US-09-157-021-35	Sequence 35, Appl	830	15	0.7	7808	3	US-08-949-386-22	Sequence 22, Appl

C 831	15	0.7	7808	3	US-08-450-562-22	Sequence 22, Appl	C 904	15	0.7	21126	1	US-08-458-831-19	Sequence 19, Appl
C 832	15	0.7	7808	4	US-08-984-709A-22	Sequence 22, Appl	C 905	15	0.7	23071	4	US-09-262-773-210	Sequence 210, App
C 833	15	0.7	7808	4	US-08-450-272-22	Sequence 22, Appl	C 906	15	0.7	28701	4	US-09-341-587-7	Sequence 7, Appl
C 834	15	0.7	7904	1	US-08-316-239B-2	Sequence 1, Appl	C 907	15	0.7	31728	4	US-09-453-702B-64	Sequence 64, Appl
C 835	15	0.7	7904	1	US-08-316-239B-2	Sequence 2, Appl	C 908	15	0.7	38844	4	US-09-734-675-5	Sequence 3, Appl
C 836	15	0.7	7904	1	US-08-410-005-1	Sequence 1, Appl	C 909	15	0.7	45546	4	US-09-146-053-6	Sequence 6, Appl
C 837	15	0.7	7904	3	US-08-929-140-1	Sequence 1, Appl	C 910	15	0.7	45716	4	US-08-965-048-5	Sequence 5, Appl
C 838	15	0.7	7904	3	US-09-560-579A-1	Sequence 1, Appl	C 911	15	0.7	45989	4	US-08-965-048-6	Sequence 6, Appl
C 839	15	0.7	7941	4	US-09-816-703A-1	Sequence 1, Appl	C 912	15	0.7	49136	4	US-09-422-869-1	Sequence 1, Appl
C 840	15	0.7	8012	4	US-09-182-117-1	Sequence 1, Appl	C 913	15	0.7	50000	4	US-09-146-053-3	Sequence 3, Appl
C 841	15	0.7	8031	4	US-09-643-597-254	Sequence 254, App	C 914	15	0.7	50000	4	US-09-146-053-3	Sequence 3, Appl
C 842	15	0.7	8050	4	US-09-491-362-11	Sequence 11, Appl	C 915	15	0.7	50000	4	US-09-146-053-3	Sequence 3, Appl
C 843	15	0.7	8050	4	US-09-874-562-11	Sequence 11, Appl	C 916	15	0.7	51259	3	US-08-781-891-209	Sequence 209, App
C 844	15	0.7	8220	2	US-08-568-459A-11	Sequence 11, Appl	C 917	15	0.7	51952	3	US-08-947-823-1	Sequence 1, Appl
C 845	15	0.7	8220	2	US-08-487-826B-11	Sequence 11, Appl	C 918	15	0.7	59065	4	US-09-813-817-3	Sequence 3, Appl
C 846	15	0.7	8220	4	US-09-210-288-11	Sequence 11, Appl	C 919	15	0.7	59065	4	US-09-813-817-3	Sequence 3, Appl
C 847	15	0.7	8418	4	US-09-182-117-5	Sequence 5, Appl	C 920	15	0.7	70000	4	US-09-851-896-3	Sequence 3, Appl
C 848	15	0.7	8651	4	US-08-961-527-181	Sequence 181, App	C 921	15	0.7	72604	4	US-09-268-992-7	Sequence 7, Appl
C 849	15	0.7	8798	4	US-09-182-117-4	Sequence 4, Appl	C 922	15	0.7	72604	4	US-09-268-992-7	Sequence 7, Appl
C 850	15	0.7	8868	3	US-08-815-809-7	Sequence 7, Appl	C 923	15	0.7	72928	3	US-09-657-474-7	Sequence 1, Appl
C 851	15	0.7	8898	4	US-08-961-527-69	Sequence 69, Appl	C 924	15	0.7	80246	4	US-09-009-913-1	Sequence 1, Appl
C 852	15	0.7	9171	2	US-08-629-001A-2	Sequence 2, Appl	C 925	15	0.7	80595	4	US-09-078-294-3	Sequence 3, Appl
C 853	15	0.7	9171	4	US-08-642-274D-2	Sequence 2, Appl	C 926	15	0.7	81001	4	US-09-750-580-1	Sequence 1, Appl
C 854	15	0.7	9171	4	US-08-952-127-2	Sequence 2, Appl	C 927	15	0.7	84495	4	US-09-797-906-3	Sequence 3, Appl
C 855	15	0.7	9171	4	US-08-952-014C-2	Sequence 2, Appl	C 928	15	0.7	87350	3	US-08-781-891-79	Sequence 79, Appl
C 856	15	0.7	9244	4	US-08-961-527-68	Sequence 68, Appl	C 929	15	0.7	87543	4	US-09-791-211-3	Sequence 3, Appl
C 857	15	0.7	9385	2	US-08-874-266-1	Sequence 1, Appl	C 930	15	0.7	99500	4	US-09-798-096-10	Sequence 10, Appl
C 858	15	0.7	9385	4	US-09-360-416-1	Sequence 1, Appl	C 931	15	0.7	112312	4	US-09-741-150-3	Sequence 3, Appl
C 859	15	0.7	9461	4	US-09-221-017B-513	Sequence 513, App	C 932	15	0.7	112312	4	US-09-128-155-16	Sequence 16, Appl
C 860	15	0.7	9515	1	US-08-920-812-13	Sequence 13, Appl	C 933	15	0.7	162450	4	US-09-345-882-1	Sequence 1, Appl
C 861	15	0.7	9515	1	US-08-920-827-13	Sequence 13, Appl	C 934	15	0.7	168575	4	US-09-426-290-1	Sequence 1, Appl
C 862	15	0.7	9515	1	US-08-921-177-13	Sequence 13, Appl	C 935	15	0.7	169998	4	US-09-676-610B-24	Sequence 24, Appl
C 863	15	0.7	9515	1	US-08-362-577C-13	Sequence 13, Appl	C 936	15	0.7	176373	3	US-09-128-155-17	Sequence 17, Appl
C 864	15	0.7	9515	2	US-08-920-828-13	Sequence 13, Appl	C 937	15	0.7	246240	2	US-08-724-394A-20	Sequence 20, Appl
C 865	15	0.7	9713	4	US-08-961-527-43	Sequence 43, Appl	C 938	15	0.7	246240	2	US-08-724-394A-21	Sequence 21, Appl
C 866	15	0.7	9870	1	US-08-508-836A-9	Sequence 9, Appl	C 939	15	0.7	246240	2	US-08-724-394A-22	Sequence 22, Appl
C 867	15	0.7	9897	1	US-08-961-527-10	Sequence 10, Appl	C 940	15	0.7	4403765	4	US-09-103-840A-2	Sequence 2, Appl
C 868	15	0.7	10299	2	US-08-477-451-1	Sequence 1, Appl	C 941	15	0.7	4411529	4	US-09-103-840A-1	Sequence 1, Appl
C 869	15	0.7	10357	4	US-08-961-527-191	Sequence 191, App	C 942	15	0.6	14	1	US-08-298-073-29	Sequence 29, Appl
C 870	15	0.7	10395	1	US-08-245-809-3	Sequence 3, Appl	C 943	15	0.6	14	1	US-08-794-153-29	Sequence 29, Appl
C 871	15	0.7	10396	1	US-08-245-809-5	Sequence 5, Appl	C 944	15	0.6	14	1	US-09-115-566-29	Sequence 29, Appl
C 872	15	0.7	10754	2	US-08-966-958-1	Sequence 1, Appl	C 945	15	0.6	18	3	US-08-181-664-16	Sequence 16, Appl
C 873	15	0.7	10754	2	US-09-215-817-1	Sequence 1, Appl	C 946	15	0.6	20	1	US-08-089-996-4	Sequence 4, Appl
C 874	15	0.7	10754	2	US-09-342-353-1	Sequence 1, Appl	C 947	15	0.6	20	2	US-08-478-178A-4	Sequence 4, Appl
C 875	15	0.7	10798	1	US-08-107-748-2	Sequence 2, Appl	C 948	15	0.6	20	2	US-08-488-177-4	Sequence 4, Appl
C 876	15	0.7	10798	5	PCT-US92-01385-2	Sequence 2, Appl	C 949	15	0.6	20	2	US-08-481-072A-4	Sequence 4, Appl
C 877	15	0.7	10846	4	US-09-096-219B-5	Sequence 5, Appl	C 950	15	0.6	20	2	US-08-684-336-4	Sequence 4, Appl
C 878	15	0.7	10900	4	US-09-098-219B-6	Sequence 6, Appl	C 951	15	0.6	20	3	US-08-481-066A-4	Sequence 4, Appl
C 879	15	0.7	10965	1	US-08-107-748-4	Sequence 4, Appl	C 952	15	0.6	20	3	US-08-578-615A-4	Sequence 4, Appl
C 880	15	0.7	10965	5	PCT-US92-01385-4	Sequence 4, Appl	C 953	15	0.6	20	4	US-09-225-749-4	Sequence 4, Appl
C 881	15	0.7	11091	4	US-09-134-001C-2243	Sequence 2243, App	C 954	15	0.6	20	4	US-08-829-637B-4	Sequence 4, Appl
C 882	15	0.7	11461	3	US-08-669-161A-29	Sequence 29, Appl	C 955	15	0.6	20	4	US-09-254-322-4	Sequence 4, Appl
C 883	15	0.7	11770	4	US-08-961-527-172	Sequence 172, App	C 956	15	0.6	20	5	PCT-US93-02213-4	Sequence 4, Appl
C 884	15	0.7	12666	4	US-08-961-527-137	Sequence 137, App	C 957	15	0.6	21	4	US-09-593-012-92	Sequence 9, Appl
C 885	15	0.7	13011	2	US-08-791-849A-14	Sequence 14, Appl	C 958	15	0.6	21	4	US-09-593-012-92	Sequence 9, Appl
C 886	15	0.7	13158	2	US-08-687-080-105	Sequence 105, App	C 959	15	0.6	21	4	US-09-593-012-186	Sequence 186, App
C 887	15	0.7	17656	4	US-09-433-579-3	Sequence 3, Appl	C 960	15	0.6	22	3	US-09-255-893-3	Sequence 2, Appl
C 888	15	0.7	18475	4	US-08-961-527-38	Sequence 38, Appl	C 961	15	0.6	24	1	US-07-741-940-94	Sequence 94, Appl
C 889	15	0.7	18609	4	US-08-943-731-1	Sequence 1, Appl	C 962	15	0.6	24	1	US-08-289-548A-94	Sequence 94, Appl
C 890	15	0.7	19011	2	US-08-310-356-36	Sequence 36, Appl	C 963	15	0.6	24	1	US-08-452-655A-94	Sequence 94, Appl
C 891	15	0.7	19124	1	US-08-487-826B-13	Sequence 13, Appl	C 964	15	0.6	24	1	US-08-452-655B-94	Sequence 94, Appl
C 892	15	0.7	19250	4	US-08-961-527-35	Sequence 35, Appl	C 965	15	0.6	24	2	US-08-117-952-596	Sequence 596, App
C 893	15	0.7	19390	4	US-08-961-527-86	Sequence 86, Appl	C 966	15	0.6	24	3	US-08-450-582-94	Sequence 94, Appl
C 894	15	0.7	19557	5	PCT-US92-06300-1	Sequence 1, Appl	C 967	15	0.6	24	4	US-08-449-731-94	Sequence 94, Appl
C 895	15	0.7	19877	2	US-08-816-155B-8	Sequence 8, Appl	C 968	15	0.6	27	4	US-08-584-040-5109	Sequence 5109, App
C 896	15	0.7	19877	3	US-09-079-587-8	Sequence 8, Appl	C 969	15	0.6	28	2	US-08-705-196-5	Sequence 5, Appl
C 897	15	0.7	20137	4	US-09-262-773-206	Sequence 206, App	C 970	15	0.6	30	3	US-08-448-619-1	Sequence 1, Appl
C 898	15	0.7	20138	4	US-09-262-773-9	Sequence 9, Appl	C 971	15	0.6	31	3	US-08-843-409-4	Sequence 4, Appl
C 899	15	0.7	21126	1	US-08-008-216-19	Sequence 19, Appl	C 972	15	0.6	31	3	US-08-843-409-5	Sequence 5, Appl
C 900	15	0.7	21126	1	US-08-008-216-19	Sequence 19, Appl	C 973	15	0.6	34	1	US-08-476-634-9	Sequence 9, Appl
C 901	15	0.7	21126	1	US-08-458-569-19	Sequence 19, Appl	C 974	15	0.6	34	1	US-08-464-518-9	Sequence 9, Appl
C 902	15	0.7	21126	1	US-08-458-569-19	Sequence 19, Appl	C 975	15	0.6	34	1	US-08-943-834-9	Sequence 9, Appl
C 903	15	0.7	21126	1	US-08-458-831-19	Sequence 19, Appl	C 976	15	0.6	38	4	US-09-482-607-2	Sequence 2, Appl

```
977 14 0.6 40 1 US-08-443-957-15 Sequence 15, Appl
978 14 0.6 42 2 US-08-471-112A-14 Sequence 14, Appl
979 14 0.6 44 2 US-08-823-516-115 Sequence 115, App
980 14 0.6 44 3 US-08-759-038-123 Sequence 123, App
981 14 0.6 44 3 US-08-758-314-123 Sequence 123, App
982 14 0.6 45 2 US-08-188-583-43 Sequence 43, Appl
983 14 0.6 45 2 US-08-680-326-68 Sequence 68, Appl
984 14 0.6 45 3 US-08-654-623-12 Sequence 12, Appl
985 14 0.6 46 4 US-09-313-221A-56 Sequence 56, Appl
986 14 0.6 47 2 US-08-198-583-48 Sequence 48, Appl
987 14 0.6 47 2 US-08-332-562A-64 Sequence 64, Appl
988 14 0.6 50 2 US-08-823-516-118 Sequence 118, App
989 14 0.6 50 3 US-08-759-038-126 Sequence 126, App
990 14 0.6 52 1 US-08-758-314-126 Sequence 126, App
991 14 0.6 52 1 US-08-418-859-45 Sequence 45, Appl
992 14 0.6 52 2 US-08-643-181-45 Sequence 45, Appl
993 14 0.6 52 4 US-09-485-737B-100 Sequence 100, App
994 14 0.6 54 4 US-09-449-218D-37 Sequence 37, Appl
995 14 0.6 55 4 US-09-485-737B-99 Sequence 99, Appl
996 14 0.6 56 2 US-08-190-199A-55 Sequence 55, Appl
997 14 0.6 56 2 US-08-190-199A-57 Sequence 57, Appl
998 14 0.6 60 3 US-08-654-623-14 Sequence 14, Appl
999 14 0.6 67 4 US-09-485-737B-97 Sequence 97, Appl
1000 14 0.6 70 1 US-07-974-251-2 Sequence 2, Appl
```

## ALIGNMENTS

RESULT 1  
US-09-058-489-90

```
; Sequence 90, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-08PA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 2319
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-90
```

Query Match 0.9%; Score 19; DB 3; Length 2319;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1830 ACATGTATCAATTGAT 1848  
Db 1598 ACATGTATCAATTGAT 1616

RESULT 2  
US-09-184-964-2  
; Sequence 2, Application US/09184964  
; Patent No. 6391574  
; GENERAL INFORMATION:  
; APPLICANT: Rine, Jasper D.  
; APPLICANT: Boyartchuk, Victor L.  
; APPLICANT: Ashby, Matthew N.  
; TITLE OF INVENTION: ARCL AND RCE1: ISOPRENYLATED CAAX  
; TITLE OF INVENTION: PROCESSING ENZYMES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:

```
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 2200 Sand Hill road, suite 100  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/184,964  
; FILING DATE: 03-NOV-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/902,774  
; FILING DATE: 30-JUL-1997  
; APPLICATION NUMBER: 60/023,491  
; FILING DATE: 07-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Suyat, Reginald J.  
; REGISTRATION NUMBER: 28,172  
; REFERENCE/DOCKET NUMBER: 09272-006004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/322-5070  
; TELEFAX: 650/854-0875  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2850 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: Genomic DNA  
US-09-184-964-2
```

Query Match 0.9%; Score 19; DB 4; Length 2850;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1674 TTTTCTTGTCAGAGAAAA 1692  
Db 919 TTTTCTTGTCAGAGAAAA 937

RESULT 3  
US-09-058-489-14  
; Sequence 14, Application US/09058489  
; Patent No. 6103886  
; GENERAL INFORMATION:  
; APPLICANT: Whitehead Institute for Biomedical Research  
; APPLICANT: Lahn, Bruce  
; APPLICANT: Page, David  
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of  
; TITLE OF INVENTION: the Y Chromosome  
; FILE REFERENCE: WHI97-08PA  
; CURRENT APPLICATION NUMBER: US/09/058,489  
; CURRENT FILING DATE: 1998-04-10  
; EARLIER APPLICATION NUMBER: 60/041,877  
; EARLIER FILING DATE: 1997-04-11  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 3408  
; TYPE: DNA  
; ORGANISM: Human  
US-09-058-489-14

Query Match 0.9%; Score 19; DB 3; Length 3408;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1829 AACATGTATCAATTGTA 1847  
|||||

Db 2388 AACATGTTATCAATTTTGA 2406

RESULT 4

US-09-058-489-17

Sequence 17, Application US/09058489

Patent No. 610386

GENERAL INFORMATION:

APPLICANT: Whitehead Institute for Biomedical Research

APPLICANT: Lahn, Bruce

APPLICANT: Page, David

TITLE OF INVENTION: Genes in the No. 610386-Recombining Region of

FILE REFERENCE: WH197-08PA

CURRENT APPLICATION NUMBER: US/09/058,489

CURRENT FILING DATE: 1998-04-10

EARLIER APPLICATION NUMBER: 60/041,877

EARLIER FILING DATE: 1997-04-11

NUMBER OF SEQ ID NOS: 91

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 17

LENGTH: 4416

TYPE: DNA

ORGANISM: Human

US-09-058-489-17

Query Match 0.9%; Score 19; DB 3; Length 4416;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1830 ACATGTTATCAATTTTGA 1848

Db 1598 ACATGTTATCAATTTTGA 1616

RESULT 5

US-09-058-489-13

Sequence 13, Application US/09058489

Patent No. 610386

GENERAL INFORMATION:

APPLICANT: Whitehead Institute for Biomedical Research

APPLICANT: Lahn, Bruce

APPLICANT: Page, David

TITLE OF INVENTION: Genes in the No. 610386-Recombining Region of

FILE REFERENCE: WH197-08PA

CURRENT APPLICATION NUMBER: US/09/058,489

CURRENT FILING DATE: 1998-04-10

EARLIER APPLICATION NUMBER: 60/041,877

EARLIER FILING DATE: 1997-04-11

NUMBER OF SEQ ID NOS: 91

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 13

LENGTH: 5322

TYPE: DNA

ORGANISM: Human

US-09-058-489-13

Query Match 0.9%; Score 19; DB 3; Length 5322;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1829 AACATGTTATCAATTTTGA 1847

Db 2388 AACATGTTATCAATTTTGA 2406

RESULT 6

US-08-651-155B-116

Sequence 116, Application US/08651155B

Patent No. 6365401

GENERAL INFORMATION:

APPLICANT: Mahan Dr., Michael J.

APPLICANT: Conner Mr., Christopher P.

APPLICANT: Hietchoff Mr., Douglas M.

TITLE OF INVENTION: METHOD AND PROBES FOR THE IDENTIFICATION

TITLE OF INVENTION: OF MICROBIAL GENES SPECIFICALLY INDUCED DURING HOST

TITLE OF INVENTION: INFECTION

NUMBER OF SEQUENCES: 255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chrisman, Bynum & Johnson, P.C.

STREET: 1900 Fifteenth Street

CITY: Boulder

STATE: CO

COUNTRY: USA

ZIP: 80302

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/651,155B

FILING DATE: 17-MAY-1996

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Petersen Mr., Steven C.

REGISTRATION NUMBER: 36,238

REFERENCE/DOCKET NUMBER: 17060.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 303/546-1300

TELEFAX: 303/449-5426

TELEX: ABAL475

INFORMATION FOR SEQ ID NO: 116:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: YES

US-08-651-155B-116

Query Match 0.8%; Score 18; DB 4; Length 339;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 CTCGACTTCATCATCAG 127

Db 184 CTCGACTTCATCATCAG 201

RESULT 7

US-09-222-575-116

Sequence 116, Application US/09222575

Patent No. 6387697

GENERAL INFORMATION:

APPLICANT: Yugiu, Jiang

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Xu, Jianshun

TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer

FILE REFERENCE: 210121.470

CURRENT APPLICATION NUMBER: US/09/222,575

CURRENT FILING DATE: 1998-12-28

NUMBER OF SEQ ID NOS: 174

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 116

LENGTH: 421

TYPE: DNA

ORGANISM: Human

US-09-222-575-116

Query Match 0.8%; Score 18; DB 4; Length 421;

Best Local Similarity 100.0%; Pred. No. 31;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1369 GGTCCAGAAATGAAGAG 1386  
|||||

Db 12 GGTCCAGAAATGAAGAG 29

## RESULT 8

US-09-318-443-3  
; Sequence 3, Application US/09318443  
; Patent No. 6197947  
; GENERAL INFORMATION:  
; APPLICANT: Hemmati-Brivanlou, Ali  
; APPLICANT: Weinstein, Daniel C.  
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4A111, AND METHODS OF USE  
; FILE REFERENCE: 600-1-211 N  
; CURRENT APPLICATION NUMBER: US/09/318,443  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 532  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-318-443-3

## Query Match

Best Local Similarity 100.0%; Pred. No. 31;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1330 GTTTGGATGAAGCTGAT 1347  
|||||

Db 346 GTTTGGATGAAGCTGAT 363

## RESULT 9

US-09-318-443-5  
; Sequence 5, Application US/09318443  
; Patent No. 6197947  
; GENERAL INFORMATION:  
; APPLICANT: Hemmati-Brivanlou, Ali  
; APPLICANT: Weinstein, Daniel C.  
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4A111, AND METHODS OF USE  
; FILE REFERENCE: 600-1-211 N  
; CURRENT APPLICATION NUMBER: US/09/318,443  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1536  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-318-443-5

## Query Match

Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1330 GTTTGGATGAAGCTGAT 1347  
|||||

Db 626 GTTTGGATGAAGCTGAT 643

## RESULT 10

US-09-318-443-7  
; Sequence 7, Application US/09318443  
; Patent No. 6197947  
; GENERAL INFORMATION:  
; APPLICANT: Hemmati-Brivanlou, Ali  
; APPLICANT: Weinstein, Daniel C.

; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4A111, AND METHODS OF USE  
; FILE REFERENCE: 600-1-211 N  
; CURRENT APPLICATION NUMBER: US/09/318,443  
; CURRENT FILING DATE: 1999-05-25  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 1682  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-318-443-7

## Query Match

Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1330 GTTTGGATGAAGCTGAT 1347  
|||||

Db 767 GTTTGGATGAAGCTGAT 784

## RESULT 11

US-09-255-392-1/C  
; Sequence 1, Application US/09255392  
; Patent No. 6214850  
; GENERAL INFORMATION:  
; APPLICANT: Evans, Ronald M.  
; APPLICANT: Forman, Barry M.  
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR  
; TITLE OF INVENTION: ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE USE THEREOF  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
; STREET: 444 South Flower Street, Suite 2000  
; CITY: Los Angeles  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/255,392  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,493  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Reiter, Stephen E.  
; REGISTRATION NUMBER: 31,192  
; REFERENCE/DOCKET NUMBER: P41 9958  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-546-1995  
; TELEFAX: 619-546-9392  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2005 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 352..1776  
US-09-255-392-1

## Query Match

Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 760 GATGAGACTCCATCTTT 777  
Db 1368 GATGAGACTCCATCTTT 1351

## RESULT 12

US-09-082-092-5  
; Sequence 5, Application US/09082092  
; Patent No. 6251628  
; GENERAL INFORMATION:  
; APPLICANT: Nakao, Atsuhito  
; APPLICANT: Moren, Anita  
; APPLICANT: Heuchel, Rainer  
; APPLICANT: Itoh, Susumu  
; APPLICANT: Afrakhte, Mozghan  
; APPLICANT: Souchevnytskyi, Serhiy  
; APPLICANT: Brodin, Gregor  
; APPLICANT: Landstrom, Marlene  
; APPLICANT: Heidlin, Nils-Erik  
; APPLICANT: Heidlin, Carl-Henrik  
; APPLICANT: ten Dijke, Peter  
; TITLE OF INVENTION: SMAD7 AND USES THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,092  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 60/047,221  
; FILING DATE: 20-MAY-1997  
; APPLICATION NUMBER: 60/060,465  
; FILING DATE: 30-SEP-1997  
; APPLICATION NUMBER: 60/075,940  
; FILING DATE: 25-FEB-1998  
; APPLICATION NUMBER: 60/077,033  
; FILING DATE: 06-MAR-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Van Amsterdam, John R.  
; REGISTRATION NUMBER: 40,212  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELETYPE:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1876 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 50..1327  
; OTHER INFORMATION:  
; US-09-082-092-5

Query Match 0.8%; Score 18; DB 4; Length 1876;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 631 TACAAAGTTTAAATGAA 648  
Db 1503 TACAAAGTTTAAATGAA 1520

## RESULT 13

US-09-291-922-19  
; Sequence 19, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hiltz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 19  
; LENGTH: 1914  
; TYPE: DNA  
; ORGANISM: Zea mays  
; US-09-291-922-19

Query Match 0.8%; Score 18; DB 4; Length 1914;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2121 TGCGTTTCTCTTCACG 2138  
Db 1445 TGCGTTTCTCTTCACG 1462

## RESULT 14

US-08-484-200-1/C  
; Sequence 1, Application US/08484200  
; Patent No. 5861274  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, RONALD M.  
; APPLICANT: FORMAN, BARRY M.  
; APPLICANT: KLEWER, STEVEN A.  
; APPLICANT: ONG, ESTELITA S.  
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID  
; TITLE OF INVENTION: SUPERFAMILY AND USES THEREFOR  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,200  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REITER, STEPHEN E.  
; REGISTRATION NUMBER: 31,192  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-546-1995  
; TELEFAX: 619-546-9392  
; INFORMATION FOR SEQ ID NO: 1:

Search completed: June 10, 2003, 14:53:24  
Job time : 155 secs

SEQUENCE CHARACTERISTICS:  
LENGTH: 2005 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 352..1776  
US-08-484-200-1

Query Match 0.8%; Score 18; DB 2; Length 2005;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 GATGAGACTCCATCTTT 777  
DB 1368 GATGAGACTCCATCTTT 1351

## RESULT 15

US-08-477-493-1/C  
Sequence 1, Application US/08477493  
Patent No. 5939442  
GENERAL INFORMATION:  
APPLICANT: Evans, Ronald M.  
APPLICANT: Forman, Barry M.  
TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,493  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter, Stephen E.  
REGISTRATION NUMBER: 31,192  
REFERENCE/DOCKET NUMBER: P41 9958  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-546-1995  
TELEFAX: 619-546-9382  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2005 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 352..1776  
US-08-477-493-1

Query Match 0.8%; Score 18; DB 2; Length 2005;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 GATGAGACTCCATCTTT 777  
DB 1368 GATGAGACTCCATCTTT 1351

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 10:40:24 ; Search time 487 Seconds  
(without alignments)  
10043.820 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 2172  
Sequence: 1 atggggagatgaagattggga.....tagatgatgagtcattgggat 2172

Scoring table:  
OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: N\_Geneseq\_101002.\*

```
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2172	100.0	2224	22	AAD06354
2	579	26.7	1006	21	AAA44726
3	370	17.0	677	22	AAS40979
4	255	11.7	591	22	AAS41534
5	255	11.7	591	22	ABL01169
6	255	11.7	591	23	ABL96630
7	220	10.1	2194	22	AA04793
8	220	10.1	2194	23	ABL97687
9	220	10.1	2195	22	AA04794

10	220	10.1	2195	23	ABL97688	Human testicular a
11	220	10.1	2197	22	AA04795	Human reproductive
12	220	10.1	2197	23	ABL97689	Human testicular a
13	60	2.8	60	24	ABN2617	Human spliced tran
14	23	1.1	2187	24	ABL99243	Mouse ischaemic co
15	21	1.0	1842	23	AAS53256	Haemophilus influe
16	21	1.0	1830121	17	AAT742063	Haemophilus influe
17	20	0.9	20	22	AAD06359	RT-PCR primer #2,
18	20	0.9	2320	21	AAC3735	Arabidopsis thalia
19	20	0.9	9832	20	ABL32657	Human immune syste
20	20	0.9	2344528	24	ABA03041	Listeria monocytog
21	19	0.9	28	22	AAD06357	PCR primer #2, to
22	19	0.9	138	22	ABA2946	Human foetal liver
23	19	0.9	138	22	AAK21377	Human brain expres
24	19	0.9	138	22	AAK47539	Human bone marrow
25	19	0.9	138	22	AA153370	Probe #22056 used
26	19	0.9	138	24	ABE21638	Human genome-deriv
27	19	0.9	592	22	ABA60407	Human foetal liver
28	19	0.9	592	22	AAK08685	Human brain expres
29	19	0.9	592	22	AAK4573	Human bone marrow
30	19	0.9	592	22	AA140291	Probe #8977 used t
31	19	0.9	592	24	AB509206	Human genome-deriv
32	19	0.9	760	22	AAH03354	Human CDNA clone (
33	19	0.9	811	23	ABL16909	Drosophila melanog
34	19	0.9	851	22	AAH05178	Human CDNA clone (
35	19	0.9	866	22	AAH05374	Human CDNA clone (
36	19	0.9	1122	22	AAH01654	Candida norvegica
37	19	0.9	1233	21	AAA29206	Human topoisomeras
38	19	0.9	1290	22	AAH00497	Pichia anomala nuc
39	19	0.9	1291	22	AAH00498	Pichia anomala nuc
40	19	0.9	1472	21	AAC98050	Human colon cancer
41	19	0.9	1472	21	AAH90759	CPE 63 coding sequ
42	19	0.9	1898	22	AAH14078	Human CDNA sequenc
43	19	0.9	1987	21	AAC77891	Human cancer assoc
44	19	0.9	2123	22	AAH13995	Human CDNA sequenc
45	19	0.9	2227	22	AAK94365	Human full-length
46	19	0.9	2262	24	ABK34739	Human CDNA for nov
47	19	0.9	2293	22	AAH15834	Human CDNA sequenc
48	19	0.9	2306	22	AAH17861	Human CDNA sequenc
49	19	0.9	2487	22	AAH15763	Human CDNA sequenc
50	19	0.9	2592	23	AAE68253	DNA encoding novel
51	19	0.9	2722	24	AAE4414	Human RNA metacoll
52	19	0.9	2850	19	AAV17602	Yeast RCE1 gene en
53	19	0.9	3039	23	AAE68807	DNA encoding novel
54	19	0.9	3039	23	AAE89721	DNA encoding novel
55	19	0.9	3041	22	AAH17625	Human CDNA sequenc
56	19	0.9	3123	22	AAH14548	Human CDNA sequenc
57	19	0.9	3408	19	AAV69632	Dead Box X (DBX) g
58	19	0.9	3550	19	AAV44295	Human secreted pro
59	19	0.9	3550	22	AAE98469	Human CDNA clone C
60	19	0.9	3764	23	ABL16908	Drosophila melanog
61	19	0.9	3862	21	AAD01223	Human semaphorin 6
62	19	0.9	4416	19	AAV69633	Dead Box Y (DBY) g
63	19	0.9	5321	21	AAE29207	Human DBX1 coding
64	19	0.9	5322	19	AAV69631	Dead Box X (DBX) g
65	19	0.9	5526	23	AAE76845	DNA encoding novel
66	19	0.9	5526	23	AAE76845	Human polynucleoti
67	19	0.9	5804	22	AAK52102	Human polynucleoti
68	19	0.9	6060	22	AAH18729	Human CDNA sequenc
69	19	0.9	10753	22	AAE85975	Domestic Pig endog
70	19	0.9	26781	23	ABL01850	Drosophila melanog
71	19	0.9	35498	23	ABL20316	Drosophila melanog
72	19	0.9	50000	24	ABA98944	Human asthma-assoc
73	19	0.9	534720	19	AAV30458	Rhizobium species
74	19	0.9	536165	19	AAV30459	Rhizobium species
75	19	0.8	98	21	AAC28944	Human secreted pro
76	18	0.8	98	21	AAZ42352	Human 5' EST isolat
77	18	0.8	108	23	ABV53358	Human prostate exp
78	18	0.8	174	22	AAE88278	H. tuberculara Hth
79	18	0.8	333	21	AAC28563	Human secreted pro
80	18	0.8	358	22	AAE24860	Human breast cance
81	18	0.8	379	21	AAC94751	Cat flea hindgut a
82	18	0.8	381	21	AAC94230	Cat flea head and

83	18	0.8	421	22	AA5A7061	Human breast cancer	156	18	0.8	13467	22	AA105944	Human reproductive
84	18	0.8	421	22	AAFI7631	Human breast cancer	157	18	0.8	13467	22	AA527670	DNA encoding novel
85	18	0.8	421	24	ABK95096	Human breast tumor	158	18	0.8	13467	22	AA527838	DNA encoding novel
86	18	0.8	423	22	ABA58227	Human foetal liver	159	18	0.8	13467	22	AAK85539	Human immune/haema
87	18	0.8	423	22	AAK06314	Human brain expres	160	18	0.8	15558	22	ABL98508	Human testicular/haema
88	18	0.8	423	22	AAK31981	Human bone marrow	161	18	0.8	15558	22	AA539926	Genomic sequence #
89	18	0.8	423	22	AAI37842	Probe #6528 used t	162	18	0.8	15558	22	AAK90370	Human digestive sy
90	18	0.8	423	24	ABS06745	Human genome-deriv	163	18	0.8	18385	22	AA528584	Genomic sequence #
91	18	0.8	452	22	AAH01646	Aspergillus fumiga	164	18	0.8	18385	22	AAK79804	Human immune/haema
92	18	0.8	517	24	ABK94752	Neurodegenerative	165	18	0.8	19399	23	ABL06008	Drosophila melanog
93	18	0.8	532	22	AAAD20194	Human translation	166	18	0.8	20020	22	AA529216	Genomic sequence #
94	18	0.8	532	22	AAAF6070	Human translation	167	18	0.8	20020	22	AA105535	Human reproductive
95	18	0.8	535	22	AAAC6920	Arabidopsis thalia	168	18	0.8	349980	22	AAH41225	Pyrococcus abyssi
96	18	0.8	587	21	AAK43690	Arabidopsis thalia	169	18	0.8	349980	22	AAH41225	Pyrococcus abyssi
97	18	0.8	599	24	ABL63546	Breast cancer rela	170	18	0.8	1664976	15	AAV21209	Methanococcus jann
98	18	0.8	733	20	AAZ15038	Human gene expres	171	17	0.8	24	24	ABQ00895	Oligonucleotide ad
99	18	0.8	777	24	ABO55513	Human ovarian anti	172	17	0.8	24	24	ABQ005732	Oligonucleotide ad
100	18	0.8	800	22	AAK63791	Human immune/haema	173	17	0.8	24	24	ABQ005773	Oligonucleotide ad
101	18	0.8	924	22	AA573663	DNA encoding novel	174	17	0.8	37	23	ABK05404	Human NOGO G-Cleav
102	18	0.8	929	22	ABAF77019	Proliferative glom	175	17	0.8	135	22	ABAF74956	Human foetal liver
103	18	0.8	957	22	AAH31733	Human olfactory re	176	17	0.8	135	22	AAK49601	Human bone marrow
104	18	0.8	1018	22	AAH00689	Cunninghamella ber	177	17	0.8	135	22	AAI26709	Probe #16642 for g
105	18	0.8	1150	21	AAAC42061	Arabidopsis thalia	178	17	0.8	135	22	AAI55483	Probe #24169 used
106	18	0.8	1205	24	ABK65178	Arabidopsis cDNA e	179	17	0.8	135	24	ABS23086	Human genome-deriv
107	18	0.8	1277	22	AAH00470	Candida dubliniens	180	17	0.8	155	22	ABA71193	Human foetal liver
108	18	0.8	1289	23	ABL06963	Drosophila melanog	181	17	0.8	155	22	AAK14991	Human brain expres
109	18	0.8	1388	23	ABL29731	Drosophila melanog	182	17	0.8	155	22	AAK45482	Human bone marrow
110	18	0.8	1536	22	AAAD20195	Human translation	183	17	0.8	155	22	AAI51427	Probe #20113 used
111	18	0.8	1536	22	AAAF66071	Human translation	184	17	0.8	155	24	ABS19750	Human genome-deriv
112	18	0.8	1558	23	ABL23041	Drosophila melanog	185	17	0.8	168	22	ABA49389	Human breast cell
113	18	0.8	1576	22	AA544681	Human full-length	186	17	0.8	168	22	ABA67299	Human foetal liver
114	18	0.8	1682	22	AAAD20196	Human translation	187	17	0.8	168	22	ABA34392	Probe #12858 for g
115	18	0.8	1682	22	AAAF66072	Human translation	188	17	0.8	168	22	AAK15736	Human brain expres
116	18	0.8	1682	24									

229	17	0.8	471	22	AA112321	Probe #2254 for ge	C 302	17	0.8	828	22	AAK90401	Human digestive sy
230	17	0.8	471	22	AA133678	Probe #2364 used t	303	17	0.8	888	23	AAK93663	DNA encoding novel
231	17	0.8	471	22	AA102237	Probe #2228 used t	304	17	0.8	909	23	AAK62810	Shrimp white spot
232	17	0.8	471	22	ABSO2194	Human genome-deriv	C 305	17	0.8	915	23	AAI16542	Human novel protei
C 233	17	0.8	472	22	ABA502030	Human foetal liver	C 306	17	0.8	960	23	AAK57149	DNA encoding novel
C 234	17	0.8	472	22	ABA56987	Human foetal liver	307	17	0.8	987	24	ABK74561	Bacillus lichenifo
C 235	17	0.8	472	22	ABA21844	Probe #310 for gen	C 308	17	0.8	1038	21	AAK46376	Arabidopsis thalia
C 236	17	0.8	472	22	AAK00314	Human brain expres	C 309	17	0.8	1041	21	AAK49125	Arabidopsis thalia
C 237	17	0.8	472	22	AAK25756	Human bone marrow	310	17	0.8	1053	22	AAK32438	Human ribonucleot
C 238	17	0.8	472	22	AAI10385	Probe #318 for gen	311	17	0.8	1053	22	AAK32447	Human ribonucleot
C 239	17	0.8	472	22	AAI15216	Probe #319 for ge	312	17	0.8	1071	23	AAK91817	DNA encoding novel
C 240	17	0.8	472	22	AAI13637	Probe #223 used to	313	17	0.8	1081	22	AAK32440	Human ribonucleot
C 241	17	0.8	472	22	AAI00320	Probe #311 used to	C 314	17	0.8	1089	21	AAA30581	Human G protein-co
C 242	17	0.8	472	24	ABSO0336	Human genome-deriv	C 315	17	0.8	1089	21	AAA30710	DNA encoding human
C 243	17	0.8	476	21	AAK35898	Arabidopsis thalia	C 316	17	0.8	1089	24	AAK26882	Human G-protein co
244	17	0.8	476	22	AAK06895	Human foetal liver	C 317	17	0.8	1125	21	AAK43034	Arabidopsis thalia
245	17	0.8	476	22	AAK06895	Human brain expres	318	17	0.8	1134	7	AAK60428	Sequence encoding
246	17	0.8	476	22	AAK32548	Human bone marrow	319	17	0.8	1182	22	AAK65738	C glutaminc codin
247	17	0.8	476	22	AAI18374	Probe #7060 used t	C 320	17	0.8	1196	14	AAQ34594	Clone for TGF-beta
248	17	0.8	476	24	ABSO7338	Human genome-deriv	321	17	0.8	1206	22	AAK66931	C glutaminc codin
249	17	0.8	483	18	AAV24801	H. pylori ORF 14cp	322	17	0.8	1237	8	AAK70314	Sequence encoding
C 250	17	0.8	492	24	ABL36418	Human colon tumour	323	17	0.8	1239	21	AAK43222	Arabidopsis thalia
C 251	17	0.8	502	21	AAK44406	Arabidopsis thalia	324	17	0.8	1266	22	AAK71669	Cornebacterium gl
C 252	17	0.8	512	22	AAK62431	Human foetal liver	325	17	0.8	1305	21	AAK46486	Arabidopsis thalia
C 253	17	0.8	512	22	AAK36647	Human bone marrow	326	17	0.8	1338	9	AAK80040	Sequence encoding
C 254	17	0.8	512	22	AAI17500	Probe #7433 for ge	C 327	17	0.8	1362	23	AAK66186	DNA encoding novel
C 255	17	0.8	512	22	AAI142410	Probe #11096 used	C 328	17	0.8	1375	23	AAK24711	Human prostate exp
C 256	17	0.8	512	24	ABN51645	Human genome-deriv	329	17	0.8	1381	21	AAK36272	Arabidopsis thalia
C 257	17	0.8	517	24	ABN51631	Human cancer relat	330	17	0.8	1389	15	AAK07077	Arabidopsis thalia
C 258	17	0.8	532	24	ABN53072	Human prostate relat	C 331	17	0.8	1389	15	AAK07077	TATA-binding prote
C 259	17	0.8	538	23	ABV08756	Human prostate exp	332	17	0.8	1401	15	AAQ72189	Rat PACAP receptor
C 260	17	0.8	541	24	ABO40246	Oligonucleotide fo	C 333	17	0.8	1446	22	AAI15870	Muskmelon 9-HPL CD
261	17	0.8	541	24	ABO40247	Oligonucleotide fo	334	17	0.8	1453	18	AAK617924	Streptococcus pneu
C 262	17	0.8	554	21	AAK40229	Arabidopsis thalia	335	17	0.8	1464	24	ABO67922	Listeria monocytog
263	17	0.8	555	18	AAV24914	H. pylori secreted	336	17	0.8	1464	24	ABO69939	Listeria monocytog
264	17	0.8	574	22	AAH10046	Human cDNA clone	C 337	17	0.8	1485	15	AAQ72191	Rat PACAP receptor
265	17	0.8	574	22	AAK93944	Primer specific fo	338	17	0.8	1517	20	AAK220021	Soybean britle-1
266	17	0.8	605	21	AAK14197	Human colon tumour	339	17	0.8	1525	21	AAK49500	Arabidopsis thalia
267	17	0.8	624	23	ABV53546	Aspergillus oryzae	340	17	0.8	1527	23	AAK90226	DNA encoding novel
C 268	17	0.8	624	21	AAK08843	Human prostate exp	341	17	0.8	1528	21	AAK33857	Arabidopsis thalia
C 269	17	0.8	626	23	ABV38653	Human prostate exp	C 342	17	0.8	1539	21	AAK46486	Arabidopsis thalia
C 270	17	0.8	626	23	ABV38653	Human prostate exp	343	17	0.8	1548	21	AAK40278	Arabidopsis thalia
C 271	17	0.8	632	23	AAK71047	DNA encoding novel	C 344	17	0.8	1560	21	AAK31381	Arabidopsis thalia
C 272	17	0.8	632	21	AAK08951	Fusarium venenat	C 345	17	0.8	1575	21	AAK39261	Arabidopsis thalia
C 273	17	0.8	663	21	AAK54144	Neisseria gonorrhe	C 346	17	0.8	1578	17	AAK42223	Human TATA-binding
274	17	0.8	697	20	AAK30352	DNA encoding a hum	C 347	17	0.8	1578	18	AAK79597	TATA-binding prote
C 275	17	0.8	702	19	AAV07909	Human cytochrome-r	C 348	17	0.8	1578	18	ABU04223	Drosophila melanog
C 276	17	0.8	702	21	AAK5134	CDNA encoding a pa	C 349	17	0.8	1601	22	AAH14924	Human cDNA sequenc
C 277	17	0.8	706	21	AAK36936	Arabidopsis thalia	C 350	17	0.8	1602	21	AAK51388	Arabidopsis thalia
C 278	17	0.8	719	21	AAK45206	Arabidopsis thalia	C 351	17	0.8	1617	19	AAV21658	AAV4 VP3 coat prot
C 279	17	0.8	720	22	ABK07921	Human ovarian and	352	17	0.8	1618	21	AAK94725	Soybean auxin tran
C 280	17	0.8	720	22	AAK03738	Human reproductive	353	17	0.8	1640	23	AAK90225	DNA encoding novel
C 281	17	0.8	720	22	AAK4566	Human immune/haema	C 354	17	0.8	1648	22	AAH13852	Human cDNA sequenc
282	17	0.8	722	21	AAV8467	Plant SDF polynuci	355	17	0.8	1666	21	AAK48754	Arabidopsis thalia
283	17	0.8	744	23	AAK55726	Streptococcus pneu	356	17	0.8	1670	21	AAK31813	Arabidopsis thalia
284	17	0.8	744	23	AAK55887	Streptococcus pneu	357	17	0.8	1674	22	AAK06174	Human reproductive
C 285	17	0.8	765	24	ABN66076	Streptococcus poly	358	17	0.8	1674	22	AAK06177	Human reproductive
C 286	17	0.8	765	24	ABN66076	Streptococcus poly	359	17	0.8	1674	22	ABK98739	Human testicular a
C 287	17	0.8	768	20	AAK56830	Human pdkk-2 cDNA	360	17	0.8	1674	23	ABK98742	Human testicular a
C 288	17	0.8	769	18	AAK55526	Human neuroblastom	C 361	17	0.8	1700	23	AAK585309	DNA encoding novel
C 289	17	0.8	778	18	AAK30821	Streptococcus pneu	C 362	17	0.8	1711	21	AAK47230	Arabidopsis thalia
C 290	17	0.8	780	21	AAK08842	Human DKK-2 DNA	C 363	17	0.8	1715	22	AAK15894	Muskmelon 9-HPL CD
C 291	17	0.8	789	21	AAK75139	Open reading frame	C 364	17	0.8	1757	21	AAK41794	Arabidopsis thalia
292	17	0.8	795	21	AAK34416	Arabidopsis thalia	C 365	17	0.8	1800	19	AAK21657	AAV4 VP2 coat prot
293	17	0.8	801	21	AAK07750	Fusarium venenat	C 366	17	0.8	1802	23	ABK20783	Drosophila melanog
C 294	17	0.8	820	22	AAK07639	Human cDNA clone	C 367	17	0.8	1813	21	AAK464012	DNA encoding a 9-h
C 295	17	0.8	828	22	AAK39954	Genomic sequence #	C 368	17	0.8	1813	21	AAK49782	Cucumber hydropero
C 296	17	0.8	828	22	AAK39955	Genomic sequence #	369	17	0.8	1846	23	AAK56381	DNA encoding novel
C 297	17	0.8	828	22	AAK39956	Genomic sequence #	370	17	0.8	1869	15	AAK072205	Rat PACAP receptor
C 298	17	0.8	828	22	AAK39957	Genomic sequence #	371	17	0.8	1879	22	AAK50386	Rat caeapase recepr
C 299	17	0.8	828	22	AAK30398	Human digestive sy	372	17	0.8	1894	22	ABK21440	Human nervous syst
C 300	17	0.8	828	22	AAK90399	Human digestive sy	373	17	0.8	1896	22	ABK21439	Human nervous syst
C 301	17	0.8	828	22	AAK90400	Human digestive sy	C 374	17	0.8	1906	19	AAK39012	CD30 ligand gene u

C 375	17	0.8	1906	24	ABK83816	Human cDNA differe	C 448	17	0.8	4494	22	AAD18129	Rat TGF-beta Resis
C 376	17	0.8	1924	19	AAV27403	Streptococcus pneu	C 449	17	0.8	4512	22	AAH16662	Human cDNA sequenc
C 377	17	0.8	1924	24	ABO84871	S. pneumoniae Sp09	C 450	17	0.8	4746	22	ABJ23508	Drosophila melanog
C 378	17	0.8	1926	21	AAC36370	Arabidopsis thalia	C 451	17	0.8	4757	19	AAV21648	AAV genome. Aden
C 379	17	0.8	1929	18	AAAT72642	Human testis-speci	C 452	17	0.8	4948	22	AAK84811	Human immune/haema
C 380	17	0.8	1929	20	AAAX09015	Testis specific is	C 453	17	0.8	4955	22	AAK32439	Human ribonucleoti
C 381	17	0.8	1930	21	AAAC45681	Arabidopsis thalia	C 454	17	0.8	4998	22	AAAS01079	Mouse vitronectin
C 382	17	0.8	2028	22	AAAF5106	S. tuberosum CBP4	C 455	17	0.8	4998	22	AAAS01079	Mouse vitronectin
C 383	17	0.8	2060	23	ABLO4259	Drosophila melanog	C 456	17	0.8	5021	22	AAK81320	Human immune/haema
C 384	17	0.8	2120	21	AAZ56978	Arabidopsis E3-2.8	C 457	17	0.8	5220	23	ABJ28025	Drosophila melanog
C 385	17	0.8	2120	21	AAZ56978	DNA encoding novel	C 458	17	0.8	5489	20	ABJ34120	Human immune syste
C 386	17	0.8	2127	23	AAAS8963	DNA encoding a S.	C 459	17	0.8	5850	20	AAK83476	Group B streptococ
C 387	17	0.8	2179	19	AAAT98571	AAV4 VP1 capsid pr	C 460	17	0.8	5881	22	AAAL07230	Human reproductive
C 388	17	0.8	2219	19	AAAD33878	Human m2404 cDNA.	C 461	17	0.8	5881	22	AAAL07230	Human reproductive
C 389	17	0.8	2219	23	ABLO9662	Drosophila melanog	C 462	17	0.8	5930	24	ABJ33394	Human immune syste
C 390	17	0.8	2235	21	AAAO5407	Streptococcus pneu	C 463	17	0.8	6051	22	AAAS46604	Tumour suppressor
C 391	17	0.8	2256	24	ABN67495	Streptococcus poly	C 464	17	0.8	6051	22	ABJ33951	Human immune syste
C 392	17	0.8	2256	24	ABN70389	Streptococcus poly	C 465	17	0.8	6135	23	ABJ33953	Human prostate exp
C 393	17	0.8	2291	22	AAI59445	Human polynucleoti	C 466	17	0.8	6135	23	ABJ24116	Human prostate exp
C 394	17	0.8	2294	21	AAAC63433	Human secreted pro	C 467	17	0.8	6135	23	ABJ24962	Human prostate exp
C 395	17	0.8	2321	22	AAI58680	Human polynucleoti	C 468	17	0.8	6135	23	ABJ25623	Human prostate exp
C 396	17	0.8	2321	22	AAI58680	S. pneumoniae derl	C 469	17	0.8	6135	23	ABJ299836	Human prostate exp
C 397	17	0.8	2325	19	AAZ56293	Human cDNA sequenc	C 470	17	0.8	6135	23	ABJ299836	Human prostate exp
C 398	17	0.8	2326	22	AAH18611	Arabidopsis thalia	C 471	17	0.8	6226	22	AAAT98451	Human ALK gene. H
C 399	17	0.8	2344	21	AAAC38986	Human polynucleoti	C 472	17	0.8	6582	18	AAAT98451	S. carnosus nitrat
C 400	17	0.8	2346	22	AAAC37726	Human polynucleoti	C 473	17	0.8	6582	18	AAAT98451	Drosophila melanog
C 401	17	0.8	2366	22	AAI60510	Claesporium oxyss	C 474	17	0.8	6853	23	ABJ04258	Drosophila melanog
C 402	17	0.8	2386	19	AAV35645	Human huntingtin 1	C 475	17	0.8	7215	21	AAAD00831	Hybrid adeno-asso
C 403	17	0.8	2510	22	AAH75515	Rat PACAP receptor	C 476	17	0.8	7228	21	ABJ04258	Drosophila melanog
C 404	17	0.8	2548	15	AAQ72206	DNA encoding novel	C 477	17	0.8	8151	21	AAAD00831	Chimeric adeno-ss
C 405	17	0.8	2559	23	AAAS78444	DNA encoding novel	C 478	17	0.8	8151	21	AAV99559	Arabidopsis lysine
C 406	17	0.8	2559	23	AAAS87347	DNA encoding novel	C 479	17	0.8	8197	24	ABJ70541	Chemically treated
C 407	17	0.8	2559	23	AAAS89198	DNA encoding novel	C 480	17	0.8	8197	24	ABJ34514	Human metastasis a
C 408	17	0.8	2560	23	AAAS98227	Human tumour marke	C 481	17	0.8	8440	23	ABJ28024	Drosophila melanog
C 409	17	0.8	2564	24	ABJ58957	cDNA encoding a hu	C 482	17	0.8	8512	19	AAV52262	Streptococcus pneu
C 410	17	0.8	2596	21	AAA12411	Partial ALK gene.	C 483	17	0.8	9453	21	AAAT21116	Human low adenosin
C 411	17	0.8	2608	16	AAQ95541	Alppha-DNA polymera	C 484	17	0.8	9453	21	AAA34994	Human cDNA differe
C 412	17	0.8	2619	15	AAQ55724	Fungal ZBC gene se	C 485	17	0.8	9453	24	ABK84194	Lung cancer relate
C 413	17	0.8	2715	24	ABN79809	Nucleotide sequenc	C 486	17	0.8	9453	24	ABJ65014	Human reproductive
C 414	17	0.8	2736	22	AAH42607	Homo sapiens vascu	C 487	17	0.8	10198	22	AAJ06056	Human testicular a
C 415	17	0.8	2846	19	AAAV20806	Human VEGF-D codin	C 488	17	0.8	10198	22	ABJ06056	Human testicular a
C 416	17	0.8	2846	22	AAA91006	Human secreted pro	C 489	17	0.8	10198	22	ABJ06056	Human testicular a
C 417	17	0.8	2894	21	AAAC63431	Drosophila melanog	C 490	17	0.8	10366	22	ABJ198621	Human nervous syst
C 418	17	0.8	2927	23	ABJ13202	Human G-protein co	C 491	17	0.8	10366	22	ABJ198621	Human nervous syst
C 419	17	0.8	2932	24	AAAD26831	Human G-protein co	C 492	17	0.8	10514	23	ABJ28624	Chemically treated
C 420	17	0.8	2932	24	AAAD26831	Human G-protein co	C 493	17	0.8	10872	23	ABJ70335	Human gene regulat
C 421	17	0.8	2986	21	AAA23448	cDNA encoding huma	C 494	17	0.8	10872	24	AA661288	Human gene regulat
C 422	17	0.8	3112	15	AAQ53948	Thermophilic DNA p	C 495	17	0.8	10872	24	ABK31388	Signal transductio
C 423	17	0.8	3115	23	ABJ17416	Drosophila melanog	C 496	17	0.8	10910	23	ABJ03131	Drosophila melanog
C 424	17	0.8	3117	22	AAH24264	Human SCP-1 coding	C 497	17	0.8	11389	19	AAV34334	Human glycoengin-2
C 425	17	0.8	3124	22	AAAF4826	Reps-like protein	C 498	17	0.8	11735	22	AAAS45329	Chemically pretrea
C 426	17	0.8	3170	21	AAA37042	Human PRO1316 (UNQ	C 499	17	0.8	11735	24	AA661142	Human gene regulat
C 427	17	0.8	3170	22	AAAF46049	Human DNA encoding	C 500	17	0.8	11735	24	ABK28168	DNA transcription
C 428	17	0.8	3170	22	AAAF46049	DNA encoding prote	C 501	17	0.8	12025	24	ABJ33299	Human immune syste
C 429	17	0.8	3393	20	AAAX24681	Human synaptonemal	C 502	17	0.8	12831	22	AAK70373	Human immune/haema
C 430	17	0.8	3393	22	AAH76512	Human SCP-1 coding	C 503	17	0.8	13165	22	AAAO7140	Human digestive c
C 431	17	0.8	3422	22	AAAL03358	Human reproductive	C 504	17	0.8	13165	22	AAK89348	Human digestive sy
C 432	17	0.8	3422	22	AAAL03360	Human immunoglobul	C 505	17	0.8	14119	23	ABJ17974	Drosophila melanog
C 433	17	0.8	3422	22	AAAS28909	Human immunoglobul	C 506	17	0.8	14800	24	ABJ66291	Lung cancer relate
C 434	17	0.8	3422	22	AAAS28911	Human cDNA sequenc	C 507	17	0.8	14835	24	AAAS94858	Human DNA sequenc
C 435	17	0.8	3551	22	AAH18414	Human cDNA sequenc	C 508	17	0.8	15240	21	AAAT21117	Human low adenosin
C 436	17	0.8	3587	23	ABLO4840	Drosophila melanog	C 509	17	0.8	15240	21	AAAT21117	Human low adenosin
C 437	17	0.8	3696	21	AAA75138	cDNA encoding a hu	C 510	17	0.8	15387	24	ABJ32184	Human immune syste
C 438	17	0.8	3718	24	ABK63564	Rat sequence diffe	C 511	17	0.8	16838	24	AAAS16667	DNA encoding super
C 439	17	0.8	3748	23	ABLO6888	Drosophila melanog	C 512	17	0.8	17512	23	ABJ09034	Drosophila melanog
C 440	17	0.8	3752	19	AAV52341	Streptococcus pneu	C 513	17	0.8	17594	24	ABJ34026	Human immune syste
C 441	17	0.8	4066	19	AAV34331	Human glycoengin-2	C 514	17	0.8	18234	22	AAH62719	Shrimp white spot
C 442	17	0.8	4066	19	AAV34331	Human glycoengin-2	C 515	17	0.8	18574	23	ABJ09756	Drosophila melanog
C 443	17	0.8	4285	22	AAI58272	Human polynucleoti	C 516	17	0.8	18574	23	ABJ17988	Drosophila melanog
C 444	17	0.8	4287	22	AAI60058	Human polynucleoti	C 517	17	0.8	18574	23	ABJ17988	Drosophila melanog
C 445	17	0.8	4359	22	AAI35981	Human musculoskele	C 518	17	0.8	18786	23	AAAS9579	Propionibacterium
C 446	17	0.8	4360	22	AAI35982	Human musculoskele	C 519	17	0.8	20813	22	ABAI7314	Human nervous syst
C 447	17	0.8	4367	23	AAAS86538	DNA encoding novel	C 520	17	0.8	28344	23	ABJ21640	Drosophila melanog

521	17	0.8	31129	24	AAD36229	Human transporter	C 594	16	0.7	236	23	AA64554	DNA encoding novel
C 522	17	0.8	32154	22	AAL37068	Human musculoskele	C 595	16	0.7	239	21	AAAC17664	Human secreted pro
C 523	17	0.8	32768	20	AAJ12954	Enterococcus faeca	C 596	16	0.7	240	21	AAAC25277	Human secreted pro
C 524	17	0.8	42324	22	AAK84724	Human immune/haema	C 597	16	0.7	244	21	AAAC08367	Human secreted pro
C 525	17	0.8	55008	22	AAK80495	Human immune/haema	C 598	16	0.7	244	21	AAAC08367	Human secreted pro
C 526	17	0.8	105184	24	ABK24122	Bacterial artificial	C 599	16	0.7	247	22	AAK76816	Human foetal liver
C 527	17	0.8	111309	24	AAK20250	Bacterial artificial	C 600	16	0.7	247	22	AAK76816	Human foetal liver
C 528	17	0.8	128600	20	ABK83461	Borrelia burgdorfe	C 601	16	0.7	247	22	AAK51447	Human brain expres
C 529	17	0.8	144460	21	AAH293815	Human cDNA differe	C 602	16	0.7	247	22	AAI57536	Human bone marrow
C 530	17	0.8	160755	23	AAH88704	Olfactory receptor	C 603	16	0.7	249	24	ABN68281	Human genome-deriv
C 531	17	0.8	305107	22	AAH62689	Human DNA sequence	C 604	16	0.7	250	22	AAK68109	Streptococcus poly
C 532	17	0.8	305107	22	AAH62689	Shrimp white spot	C 605	16	0.7	250	22	AAK68109	Human lung tumour
C 533	17	0.8	349980	22	AAH68526	C glutamylcun codin	C 606	16	0.7	252	20	AAK51981	CDNA encoding clon
C 534	17	0.8	349980	22	AAH68526	C glutamylcun codin	C 607	16	0.7	252	24	ABN26603	Human secreted pro
C 535	17	0.8	349980	22	AAH68530	C glutamylcun codin	C 608	16	0.7	254	24	ABN94848	Human ORFX polynuc
C 536	17	0.8	349980	22	AAH68533	C glutamylcun codin	C 609	16	0.7	254	24	ABN94848	Gene #1346 used to
C 537	17	0.8	513445	22	AAI61373	Soybean 318013 reg	C 610	16	0.7	259	19	AAI62614	Colon adenocarcino
C 538	17	0.8	513445	22	AAI61373	Soybean 318013 reg	C 611	16	0.7	269	19	AAI62614	Human secreted pro
C 539	17	0.8	580073	18	AAI58840	Mycoptasma genital	C 612	16	0.7	270	24	AAI61542	Lung small cell ca
C 540	17	0.8	611590	21	AAI58840	Mycoptasma genital	C 613	16	0.7	270	24	AAI61542	Lung small cell ca
C 541	17	0.8	611590	21	AAI58840	Mycoptasma genital	C 614	16	0.7	278	16	AAI26597	Human gene signatu
C 542	17	0.8	910715	20	AAK20248	Borrelia burgdorfe	C 615	16	0.7	278	21	AAK59778	Human secreted pro
C 543	17	0.8	910715	20	AAK20248	Borrelia burgdorfe	C 616	16	0.7	278	22	AAK59778	Human immune/haema
C 544	17	0.8	1230025	20	AAK91990	Listeria innocua C	C 617	16	0.7	285	24	ABU71457	Corn tassell-deriv
C 545	17	0.8	1230025	20	AAK91990	Nucleotide sequenc	C 618	16	0.7	290	24	ABU71457	Human colon cancer
C 546	17	0.8	1503900	22	AAK95240	Human neutregulin-1	C 619	16	0.7	301	24	ABN26366	Human ORFX polynuc
C 547	17	0.8	1503900	22	AAK95240	Human neutregulin-1	C 620	16	0.7	304	22	ABN26366	Human breast cell
C 548	17	0.8	1830121	17	AAK42063	Haemophilus influe	C 621	16	0.7	304	22	ABN26366	Human foetal liver
C 549	17	0.8	2155561	24	ABN71527	Streptococcus poly	C 622	16	0.7	304	22	ABN26366	Probe #44727 for g
C 550	17	0.8	2155561	24	ABN71527	Streptococcus poly	C 623	16	0.7	304	22	AAK17613	Human brain expres
C 551	17	0.8	2355589	24	ABN90521	Genomic sequence o	C 624	16	0.7	304	22	AAK3424	Human bone marrow
C 552	17	0.8	3011208	24	ABQ69245	Listeria innocua D	C 625	16	0.7	304	22	AAI24204	Probe #14137 for g
C 553	17	0.8	4403765	22	AAI99683	Mycobacterium tube	C 626	16	0.7	304	22	AAI24204	Probe #18173 used
C 554	17	0.8	4411529	22	AAI99683	Mycobacterium tube	C 627	16	0.7	304	22	AAI09766	Probe #9757 used t
C 555	16	0.7	46	21	AAH48006	Ribosomal S9 prote	C 628	16	0.7	304	24	ABN17527	Human genome-deriv
C 556	16	0.7	46	21	AAH48006	Ribosomal S9 prote	C 629	16	0.7	317	23	ABN17527	Human prostate exp
C 557	16	0.7	51	22	AAI26993	Human SNP oligonuc	C 630	16	0.7	317	23	ABN17527	Staphylococcus aur
C 558	16	0.7	60	24	ABN47942	Human spliced tran	C 631	16	0.7	327	22	AAH50173	Human polynucleoti
C 559	16	0.7	112	22	AAK20039	Human brain expres	C 632	16	0.7	330	22	AAH50173	Human prostate exp
C 560	16	0.7	112	22	AAK46083	Human bone marrow	C 633	16	0.7	333	23	ABV12865	DNA encoding Mmb 1
C 561	16	0.7	112	22	AAI51991	Probe #20677 used	C 634	16	0.7	336	15	AAO70086	Human secreted pro
C 562	16	0.7	112	24	ABN20380	Human genome-deriv	C 635	16	0.7	338	21	AAK27060	Human ovorian can
C 563	16	0.7	118	24	ABN89706	Nucleotide sequenc	C 636	16	0.7	338	16	AAI24869	Human gene signatu
C 564	16	0.7	119	21	AAAI2629	Human secreted pro	C 637	16	0.7	346	16	AAI24869	DNA encoding human
C 565	16	0.7	138	21	AAAI2629	Human secreted pro	C 638	16	0.7	350	20	AAK63145	5' cDNA sequence o
C 566	16	0.7	141	22	ABN75997	Human foetal liver	C 639	16	0.7	350	21	AAK63145	Human breast tumou
C 567	16	0.7	141	22	ABN75997	Human foetal liver	C 640	16	0.7	351	18	AAK77280	Staphylococcus aur
C 568	16	0.7	141	22	AAK24674	Human brain expres	C 641	16	0.7	351	22	AAH12793	Human cDNA clone
C 569	16	0.7	141	22	AAK50673	Human bone marrow	C 642	16	0.7	353	21	AAK50673	Human immune/haema
C 570	16	0.7	141	22	AAI27683	Probe #17616 for g	C 643	16	0.7	353	22	AAK50673	Human immune/haema
C 571	16	0.7	141	22	AAI27683	Probe #17616 for g	C 644	16	0.7	353	22	AAK50673	Human prostate exp
C 572	16	0.7	141	22	AAI27683	Probe #17616 for g	C 645	16	0.7	353	22	AAI27683	Human polynucleoti
C 573	16	0.7	141	22	AAI27683	Probe #17616 for g	C 646	16	0.7	353	22	AAI27683	Human polynucleoti
C 574	16	0.7	141	22	AAI27683	Probe #17616 for g	C 647	16	0.7	353	22	AAI27683	Human polynucleoti
C 575	16	0.7	141	22	AAI27683	Probe #17616 for g	C 648	16	0.7	353	22	AAI27683	Human polynucleoti
C 576	16	0.7	141	22	AAI27683	Probe #17616 for g	C 649	16	0.7	353	22	AAI27683	Human polynucleoti
C 577	16	0.7	141	22	AAI27683	Probe #17616 for g	C 650	16	0.7	353	22	AAI27683	Human polynucleoti
C 578	16	0.7	141	22	AAI27683	Probe #17616 for g	C 651	16	0.7	353	22	AAI27683	Human polynucleoti
C 579	16	0.7	141	22	AAI27683	Probe #17616 for g	C 652	16	0.7	353	22	AAI27683	Human polynucleoti
C 580	16	0.7	141	22	AAI27683	Probe #17616 for g	C 653	16	0.7	353	22	AAI27683	Human polynucleoti
C 581	16	0.7	141	22	AAI27683	Probe #17616 for g	C 654	16	0.7	353	22	AAI27683	Human polynucleoti
C 582	16	0.7	141	22	AAI27683	Probe #17616 for g	C 655	16	0.7	353	22	AAI27683	Human polynucleoti
C 583	16	0.7	141	22	AAI27683	Probe #17616 for g	C 656	16	0.7	353	22	AAI27683	Human polynucleoti
C 584	16	0.7	141	22	AAI27683	Probe #17616 for g	C 657	16	0.7	353	22	AAI27683	Human polynucleoti
C 585	16	0.7	141	22	AAI27683	Probe #17616 for g	C 658	16	0.7	353	22	AAI27683	Human polynucleoti
C 586	16	0.7	141	22	AAI27683	Probe #17616 for g	C 659	16	0.7	353	22	AAI27683	Human polynucleoti
C 587	16	0.7	141	22	AAI27683	Probe #17616 for g	C 660	16	0.7	353	22	AAI27683	Human polynucleoti
C 588	16	0.7	141	22	AAI27683	Probe #17616 for g	C 661	16	0.7	353	22	AAI27683	Human polynucleoti
C 589	16	0.7	141	22	AAI27683	Probe #17616 for g	C 662	16	0.7	353	22	AAI27683	Human polynucleoti
C 590	16	0.7	141	22	AAI27683	Probe #17616 for g	C 663	16	0.7	353	22	AAI27683	Human polynucleoti
C 591	16	0.7	141	22	AAI27683	Probe #17616 for g	C 664	16	0.7	353	22	AAI27683	Human polynucleoti
C 592	16	0.7	141	22	AAI27683	Probe #17616 for g	C 665	16	0.7	353	22	AAI27683	Human polynucleoti
C 593	16	0.7	141	22	AAI27683	Probe #17616 for g	C 666	16	0.7	353	22	AAI27683	Human polynucleoti

667	16	0.7	377	21	AAC53208	740	16	0.7	495	24	ABN94642	Gene #1140 used to
668	16	0.7	377	22	AA181457	741	16	0.7	496	22	AAH83682	Human ovarian tumo
669	16	0.7	381	21	AAC5971	742	16	0.7	497	22	ABA57712	Human foetal liver
670	16	0.7	381	21	AAC56750	743	16	0.7	497	22	AAK05773	Human brain expres
671	16	0.7	383	22	AAS38334	744	16	0.7	497	22	AAK13397	Human bone marrow
672	16	0.7	384	21	AAC56533	745	16	0.7	497	22	AAI137290	Probe #5976 used t
673	16	0.7	386	14	AAQ60791	746	16	0.7	497	24	ABSO6150	Human genome-deriv
674	16	0.7	387	22	AA181458	747	16	0.7	498	22	AAH19131	Interferon alpha h
675	16	0.7	387	23	ABV13412	748	16	0.7	498	22	AAH19154	Interferon alpha h
676	16	0.7	389	22	AAH64371	749	16	0.7	498	22	AAH19158	Interferon alpha h
677	16	0.7	389	24	ABL93110	750	16	0.7	498	22	AAH19160	Interferon alpha h
678	16	0.7	392	22	AAH66495	751	16	0.7	499	24	ABL16818	Human colon tumour
679	16	0.7	392	23	ABV38832	752	16	0.7	502	22	AAK11600	Human brain expres
680	16	0.7	393	24	ABK80262	753	16	0.7	505	22	AAS36292	Human cardiovascular
681	16	0.7	393	24	ABL80260	754	16	0.7	510	21	AAC37182	Arabidopsis thalia
682	16	0.7	394	20	AAK40652	755	16	0.7	510	24	ABO56308	Human colon cancer
683	16	0.7	398	21	AAC26415	756	16	0.7	512	23	ABV19333	Human prostate exp
684	16	0.7	399	20	AAK40981	757	16	0.7	512	23	ABV22747	Human prostate exp
685	16	0.7	399	20	AAK40981	758	16	0.7	512	23	ABV28575	Human prostate exp
686	16	0.7	401	22	AAK95825	759	16	0.7	513	24	ABN73788	Bovine embryonic g
687	16	0.7	401	22	AAK95326	760	16	0.7	513	21	AAC52772	Arabidopsis thalia
688	16	0.7	401	22	AAK97318	761	16	0.7	515	24	ABN61303	Human cancer relat
689	16	0.7	401	22	AAK97318	762	16	0.7	516	22	AAS24763	Human ovarian PCR-
690	16	0.7	401	22	AAK97819	763	16	0.7	516	22	AAH83396	Human ovarian tumo
691	16	0.7	403	22	AAK74255	764	16	0.7	517	24	ABK75904	Human licheniflo
692	16	0.7	403	22	AAH66702	765	16	0.7	518	22	ABAK1854	Human foetal liver
693	16	0.7	409	22	AAK61860	766	16	0.7	518	22	ABAK1854	Probe #7866 for ge
694	16	0.7	409	23	ABV12151	767	16	0.7	518	22	ABA29420	Human brain expres
695	16	0.7	409	24	ABN94473	768	16	0.7	518	22	AAK10167	Human bone marrow
696	16	0.7	410	22	AAK58995	769	16	0.7	518	22	AAK36064	Probe #7153 for ge
697	16	0.7	425	23	ABV04243	770	16	0.7	518	22	AAI17220	Human genome-deriv
698	16	0.7	427	23	ABV34530	771	16	0.7	518	24	ABSI10224	Human genome-deriv
699	16	0.7	427	23	ABV34388	772	16	0.7	522	22	AAK65802	Human immune/haema
700	16	0.7	429	21	AAK44637	773	16	0.7	523	24	ABO52910	Oligonucleotide fo
701	16	0.7	437	23	ABV33296	774	16	0.7	523	24	ABO52911	Oligonucleotide fo
702	16	0.7	438	24	ABN18644	775	16	0.7	524	24	AAS34097	Human cDNA encodin
703	16	0.7	442	23	ABV59467	776	16	0.7	528	21	AAC75522	Human ORFX ORF107
704	16	0.7	443	22	AAS32409	777	16	0.7	528	22	AAK61588	Human immune/haema
705	16	0.7	446	24	ABL79528	778	16	0.7	528	24	ABN21557	Human ORFX polynuc
706	16	0.7	446	21	AAZ80545	779	16	0.7	530	24	ABL82245	Human ovarian canc
707	16	0.7	458	23	ABV10608	780	16	0.7	533	24	ABL82317	Human ovarian canc
708	16	0.7	459	21	AAZ94512	781	16	0.7	540	22	AAH34721	Human colon cancer
709	16	0.7	461	24	ABL80648	782	16	0.7	541	22	AAS36293	Human cardiovascular
710	16	0.7	462	20	AAZ33804	783	16	0.7	544	23	ABV52539	Human prostate exp
711	16	0.7	462	22	ABL23949	784	16	0.7	547	21	AAC33806	Cat flea hindgut a
712	16	0.7	463	23	ABV47758	785	16	0.7	549	22	ABAK1658	Human foetal liver
713	16	0.7	464	21	AAK50893	786	16	0.7	549	22	AAK09959	Human brain expres
714	16	0.7	466	21	AAC37823	787	16	0.7	549	22	AAK35853	Human bone marrow
715	16	0.7	467	22	ABA58772	788	16	0.7	549	22	AAI1570	Probe #10256 used
716	16	0.7	467	22	ABA27713	789	16	0.7	549	23	ABL16109	Drosophila melanog
717	16	0.7	467	22	AAK06915	790	16	0.7	550	21	AAC52560	Arabidopsis thalia
718	16	0.7	467	22	AAK32645	791	16	0.7	555	24	ABO55631	Human ovarian anti
719	16	0.7	467	22	AAI38460	792	16	0.7	555	21	AAA45987	Human metatatic m
720	16	0.7	467	24	ABSO7436	793	16	0.7	556	22	ABA31368	Probe #9834 for ge
721	16	0.7	468	22	ABA46434	794	16	0.7	556	22	AAK12659	Human brain expres
722	16	0.7	468	24	ABSO5335	795	16	0.7	556	22	AAK38419	Human bone marrow
723	16	0.7	470	21	AAC03929	796	16	0.7	556	23	AAI19213	Probe #9146 for ge
724	16	0.7	471	22	AAH88291	797	16	0.7	566	22	ABV54235	Human prostate exp
725	16	0.7	472	22	AAI35384	798	16	0.7	566	24	ABSI2477	Human genome-deriv
726	16	0.7	472	22	AAK33541	799	16	0.7	568	22	AAK1748	Human immune/haema
727	16	0.7	472	24	ABSO8400	800	16	0.7	570	23	ABV59365	Human prostate exp
728	16	0.7	474	22	AAK07358	801	16	0.7	587	22	ABA63688	Human foetal liver
729	16	0.7	474	22	AAK33125	802	16	0.7	587	22	ABA30878	Probe #9344 for ge
730	16	0.7	474	22	AAI38921	803	16	0.7	587	22	AAK12204	Human brain expres
731	16	0.7	476	24	ABSO7943	804	16	0.7	587	22	AAK37926	Human bone marrow
732	16	0.7	476	18	AAV75780	805	16	0.7	587	22	AAI18687	Probe #6620 for ge
733	16	0.7	478	21	AAC41291	806	16	0.7	587	22	AAI43800	Human cDNA clone (
734	16	0.7	480	24	ABN63004	807	16	0.7	587	22	AAH12019	Human prostate exp
735	16	0.7	484	22	AAK54282	808	16	0.7	587	23	ABV52428	Human genome-deriv
736	16	0.7	487	22	AAK25037	809	16	0.7	587	24	ABSI1922	Human prostate exp
737	16	0.7	488	21	AAC27563	810	16	0.7	589	22	AAS23920	Human ovarian PCR-
738	16	0.7	489	23	ABV10782	811	16	0.7	589	22	AAH82478	Human ovarian tumo
739	16	0.7	492	21	AAZ43012	812	16	0.7	589	23	ABV16966	Human prostate exp







```

Db      314  TTGAAATGGTGTACTCTGGTTTCTGGAGAGAGCTAGTAATGACTGGAAATAT 373
Qy      361  CCAACACGAAACAGAGGGTTTCCAGAGAGCGGCTATCGAGATGAAATATTCAGAA 420
Db      374  CCAACACGAAACAGAGGGTTTCCAGAGAGCGGCTATCGAGATGAAATATTCAGAA 433
Qy      421  GCTTCAGGGCCATACAGAGAGGTGAAAGAGTAGTTCCAGAGTTGCCGTGAGATTT 480
Db      434  GCTTCAGGGCCATACAGAGAGGTGAAAGAGTAGTTCCAGAGTTGCCGTGAGATTT 493
Qy      481  GGTCTAGGAAGTCCAAATATGACTTGAAGCCAGACGATGTATGACGCGCATGTGGC 540
Db      494  GGTCTAGGAAGTCCAAATATGACTTGAAGCCAGACGATGTATGACGCGCATGTGGC 553
Qy      541  CTTTGTGTTCTAGAGACAGTATTAAGTGCACAGGTATGGGTACTCTCAAGC 600
Db      554  CTTTGTGTTCTAGAGACAGTATTAAGTGCACAGGTATGGGTACTCTCAAGC 613
Qy      601  AGAAGTGGCAGTGAAGTGAACGAGTGTACAAAGTTTAAATGAAGAATATAACA 660
Db      614  AGAAGTGGCAGTGAAGTGAACGAGTGTACAAAGTTTAAATGAAGAATATAACA 673
Qy      661  GGCCTTGGAAAGATTTCTTGAAGTCAAGACGAGAGAGAGAAAGTAGTATCTCAA 720
Db      674  GGCCTTGGAAAGATTTCTTGAAGTCAAGACGAGAGAGAGAAAGTAGTATCTCAA 733
Qy      721  GGACCAAAAGTGAAGTCAATACCCCTCTCCACCTGAGATGAGACTCCATCTTGGCA 780
Db      734  GGACCAAAAGTGAAGTCAATACCCCTCTCCACCTGAGATGAGACTCCATCTTGGCA 793
Qy      781  CATTATCAGACAGGCATTAACCTTCGACAAATACGACATATTCTTGTGAAGTCTGGA 840
Db      794  CATTATCAGACAGGCATTAACCTTCGACAAATACGACATATTCTTGTGAAGTCTGGA 853
Qy      841  CATATGACACGACGACATTTCTGACTTTGAAAGCTAATCTCTGTCAAGACTGTAAT 900
Db      854  CATATGACACGACGACATTTCTGACTTTGAAAGCTAATCTCTGTCAAGACTGTAAT 913
Qy      901  AACAACTGTGTAAAGCTGTATTAAGCTTACTCTCTGTCGAAAAATACAGATTCCT 960
Db      914  AACAACTGTGTAAAGCTGTATTAAGCTTACTCTCTGTCGAAAAATACAGATTCCT 973
Qy      961  ATCATACTTGCAGAGACGATTTGATGAGCTGTGCTCAACAGAGCTCTGGGAAGACTGCG 1020
Db      974  ATCATACTTGCAGAGACGATTTGATGAGCTGTGCTCAACAGAGCTCTGGGAAGACTGCG 1033
Qy      1021  GCTTTTCTCTCAACAAATTTTGGCTCATATGATGATGATGAAATTAATCTGCCAGTCTGTTT 1080
Db      1034  GCTTTTCTCTCAACAAATTTTGGCTCATATGATGATGATGAAATTAATCTGCCAGTCTGTTT 1093
Qy      1081  AAAAGATTGACGAAACGAGAGTATTAATGCTACACCAACTCGAAGATTTGCTCAACGAG 1140
Db      1094  AAAAGATTGACGAAACGAGAGTATTAATGCTACACCAACTCGAAGATTTGCTCAACGAG 1153
Qy      1141  ATTATTTTGAAGCCAGAAAAATTTTCTTTGGGACTTGTGAAGCTGTGTTATATAT 1200
Db      1154  ATTATTTTGAAGCCAGAAAAATTTTCTTTGGGACTTGTGAAGCTGTGTTATATAT 1213
Qy      1201  GGGGGAACCCAGCTGGGACATTCATTAATTCGAAATAGTACAGGCTGTATATATATAT 1260
Db      1214  GGGGGAACCCAGCTGGGACATTCATTAATTCGAAATAGTACAGGCTGTATATATATAT 1273
Qy      1261  GCTACTCTGGAAACGATGATGATATCATAGGCAAAAGAAAGATTTGCTTCAACAGATC 1320
Db      1274  GCTACTCTGGAAACGATGATGATATCATAGGCAAAAGAAAGATTTGCTTCAACAGATC 1333
Qy      1321  AAATACTAGTTTGGATGAGAGCTGATGCTGTTGGATATGGTTTGGTCCAGAAATG 1380
Db      1334  AAATACTAGTTTGGATGAGAGCTGATGCTGTTGGATATGGTTTGGTCCAGAAATG 1393
Qy      1381  AAGAAGTTAATTTCTTGCCCAAGAAATGCATCAAGGAACGCGCCAAACCTTATGTTTC 1440

```

```

Db      1394  AAGAAGTTAATTTCTTGCCCAAGAAATGCCATCAAGGAACGCGCCAAACCTTATGTTTC 1453
Qy      1441  AGTGCAACTTTTCCAGAGAAATTTCAAAGGTGGCTCAGAGTTTAAAGTCAAAATAT 1500
Db      1454  AGTGCAACTTTTCCAGAGAAATTTCAAAGGTGGCTCAGAGTTTAAAGTCAAAATAT 1513
Qy      1501  CTGTTTGTCTGTTGGACAAAGTGGTGAGCATGTAGAGATGTTACAGACCGTTCTC 1560
Db      1514  CTGTTTGTCTGTTGGACAAAGTGGTGAGCATGTAGAGATGTTACAGACCGTTCTC 1573
Qy      1561  CAAAGTGGCCAGTTCTCAAAAAGAAAGAAAGCTGTTGAATTTCTGCCAAACATAGGGGAT 1620
Db      1574  CAAAGTGGCCAGTTCTCAAAAAGAAAGAAAGCTGTTGAATTTCTGCCAAACATAGGGGAT 1633
Qy      1621  GAAAGAACTATGCTGTTGTGAACCTAAGAAAAAGCAGATTTTACTGCAACTTTCTT 1680
Db      1634  GAAAGAACTATGCTGTTGTGAACCTAAGAAAAAGCAGATTTTACTGCAACTTTCTT 1693
Qy      1681  TGTCAAGAAAAAATATCACTACAGATATCATGTGATGCGGAACAGAGAGCGGGAG 1740
Db      1694  TGTCAAGAAAAAATATCACTACAGATATCATGTGATGCGGAACAGAGAGCGGGAG 1753
Qy      1741  CAAAGCTTTTGAAGATTTTCCCTTTGGAAAGTCCCAAGTTCTTGTGCTACTTCAATGCT 1800
Db      1754  CAAAGCTTTTGAAGATTTTCCCTTTGGAAAGTCCCAAGTTCTTGTGCTACTTCAATGCT 1813
Qy      1801  GCCAGAGGGCTGATATGAATAATGTGCACATGTTATCAATTTTGAATCTTCTCTTACC 1860
Db      1814  GCCAGAGGGCTGATATGAATAATGTGCACATGTTATCAATTTTGAATCTTCTCTTACC 1873
Qy      1861  ATTGATGAATATGTTCACTACGAATTTGGCGTACTGTTGTTGGGAATACGCGAGACA 1920
Db      1874  ATTGATGAATATGTTCACTACGAATTTGGCGTACTGTTGTTGGGAATACGCGAGACA 1933
Qy      1921  ATTTCTTTTGTATCTTGAATGAGATTAACCATTTTACACAGCTCTAGTAAAGATTTG 1980
Db      1934  ATTTCTTTTGTATCTTGAATGAGATTAACCATTTTACACAGCTCTAGTAAAGATTTG 1993
Qy      1981  ACAAGTCTCAACAGATATGTTCTCGACATGTTGGAAGAAATTTGCTTTAGTACATACAT 2040
Db      1994  ACAAGTCTCAACAGATATGTTCTCGACATGTTGGAAGAAATTTGCTTTAGTACATACAT 2053
Qy      2041  CTTGGCTTCAAGTGGTGTATACAGAGAAACGTTTTCATCAGTTGATACAGAAAGGC 2100
Db      2054  CTTGGCTTCAAGTGGTGTATACAGAGAAACGTTTTCATCAGTTGATACAGAAAGGC 2113
Qy      2101  AAGAGCACTTTGAACACAGCTGGGTTTCTTCTTACAGAGCTCCCAATCCAGTAGATAT 2160
Db      2114  AAGAGCACTTTGAACACAGCTGGGTTTCTTCTTACAGAGCTCCCAATCCAGTAGATAT 2173
Qy      2161  GAGTCATGGGAT 2172
Db      2174  GAGTCATGGGAT 2185

RESULT 2
AAAA4726
ID AAA44726 standard; cDNA; 1006 BP.
AC AAA44726;
XX
XX 21-AUG-2000 (first entry)
DE Human secreted expressed sequence tag SEQ ID NO:1301.
XX
XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiasthmatic; vulnery; antiparasitomonian;
KW antitumor; osteoprotective; neuroprotective; nootropic; antiparasitic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;

```

KM insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KM lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KM central nervous system disorder; Alzheimer's disease; stroke;  
 KM Parkinson's disease; Huntington's disease; coagulation disorder;  
 KM haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KM tumour; infection; depression; psoriasis; ss.  
 XX Homo sapiens.  
 OS  
 PN WO200021991-A1.  
 XX  
 XX 20-APR-2000.  
 PD  
 XX 15-OCT-1999; 99WO-US24206.  
 PF  
 XX 15-OCT-1998; 98US-0104436.  
 PR  
 XX (GEMV ) GENETICS INST INC.  
 PA  
 XX Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR;  
 XX WPI: 2000-317938/27.  
 DR  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (ESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 XX Claim 1; Page 540; 803pp; English.  
 PS  
 XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (ESTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The ESTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antiasthmatic; vulnary; antitumor; osteopathic; neuroprotective;  
 CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The ESTs can be used for gene  
 CC therapy and in vaccines. The ESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the ESTs. Proteins encoded by the ESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumour, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA4526 to AAA4591 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 CC  
 XX  
 XX Sequence 1006 BP; 314 A; 181 C; 251 G; 260 T; 0 other;  
 SQ  
 Query Match 26 7%; Score 579; DB 21; Length 1006;  
 Best Local Similarity 99.8%; Pred. No. 5.6e-276;  
 Matches 629; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 557 CTTGGAAGTCAGAGCAGAGAGAGAGAAAGTAGTACTCAAGGCCAAAAGTGACCT 616  
 Qy 737 ACATACCCCTCCTCCACCTGAGATGAGAGCTCCATCTTTGACATTTACAGACGCGCA 796  
 Db 617 ACATACCCCTCCTCCACCTGAGATGAGAGCTCCATCTTTGACATTTACAGACGCGCA 676  
 Qy 797 TAACTTCGACAAATACGACACTATTCTTGTGAGAGTCTGACATGATGACACACGAG 856  
 Db 677 TAACTTCGACAAATACGACACTATTCTTGTGAGAGTCTGACATGATGACACACGAG 736  
 Qy 857 CAATTGACTTTTGAAGAGCTTAATCTCTGTACAGACCTGAATTAACAATTGCTTAAG 916  
 Db 737 CAATTGACTTTTGAAGAGCTTAATCTCTGTACAGACCTGAATTAACAATTGCTTAAG 796  
 Qy 917 CTGTTAATTAAGCTTACTCTGTCGCAAAATATACGATATCTATCATCTTGACAGGAC 976  
 Db 797 CTGTTAATTAAGCTTACTCTGTCGCAAAATATACGATATCTATCATCTTGACAGGAC 856  
 Qy 977 GAGATTGATGCTGTGCTCAACAGAGGTCTGGGAAGACTGGGGCTTTTCTCTACCAA 1036  
 Db 857 GAGATTGATGCTGTGCTCAACAGAGGTCTGGGAAGACTGGGGCTTTTCTCTACCAA 916  
 Qy 1037 TTTGGCTCATATGATGATGATGAATTAATCTGCCAGTCTTTTAAAGCTTGACGAGAC 1096  
 Db 917 TTTGGCTCATATGATGATGATGAATTAATCTGCCAGTCTTTTAAAGCTTGACGAGAC 976  
 Qy 1097 CAGAGTGTATTATTGTTAGCAGCAACTCGAG 1126  
 Db 977 CAGAGTGTATTATTGTTAGCAGCAACTCGAG 1006

RESULT 3  
 AAS40979  
 ID AAS40979 standard; cDNA; 677 BP.  
 XX  
 AC AAS40979;  
 XX  
 DT 17-DEC-2001 (first entry)  
 XX  
 XX cDNA encoding novel human enzyme polypeptide #195.  
 DE  
 XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
 KM ligase; hyperproliferative disorder; immunodeficiency disorder;  
 KM autoimmune disorder; neurological disorder; metabolic disorder;  
 KM inflammatory disorder; cardiovascular disorder; reproductive disorder;  
 KM blood-related disorder; infectious disorder; gene therapy; cytostatic;  
 KM anti arthritic; nephrotropic; anticoagulant; ss.  
 OS  
 XX Homo sapiens.  
 XX  
 PN WO200155301-A2.  
 PD  
 XX 02-AUG-2001.  
 PF  
 XX 17-JAN-2001; 2001WO-US01239.  
 XX  
 XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0224521.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 06-SEP-2000; 2000US-0231242.  
PR 06-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234998.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251588.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251866.  
PR 06-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254907.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465566/50.  
XX P-PSDB; AAU23109.  
XX  
XX Novel polypeptides and polynucleotides useful for diagnosing,  
PT preventing, treating neural, immune system, muscular, reproductive,  
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
PT diseases  
XX  
XX Claim 4; SEQ ID No 205; 1180bp; English.  
PS  
XX  
XX The present invention relates to the isolation of novel human enzyme  
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
CC isomerases or ligases. The sequences of the invention are useful in the  
CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
CC disorders including hyperproliferative disorders (e.g. cancer),  
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),



PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251858.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465566/50.  
DR P-PSDB; AAU23664.  
XX  
XX  
PT Novel polypeptides and polynucleotides useful for diagnosing,  
PT preventing, treating neural, immune system, muscular, reproductive,  
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
XX diseases -  
PS Claim 4; SEQ ID No 760; 1180bp; English.

XX  
CC The present invention relates to the isolation of novel human enzyme  
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
CC encoding them. The enzyme polypeptides of the invention may comprise the  
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
CC isomerases or ligases. The sequences of the invention are useful in the  
CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
CC disorders including hyperproliferative disorders (e.g. cancer),  
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
CC blood-related disorders (e.g. haemophilia), reproductive disorders  
CC (e.g. infertility) and infectious disorders (e.g. influenza). The  
CC polynucleotides of the invention can also be used in gene therapy.  
CC AAU40785-AAU41684 represent cDNA sequences encoding for the novel human  
CC enzyme polypeptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 591 BP; 203 A; 100 C; 111 G; 176 T; 1 other;  
Query Match 11.7%; Score 255; DB 22; Length 591;  
Best Local Similarity 100.0%; Pred. No. 1.3e-115;  
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1883 TTGGCGCTACTGCTCGTTGGGAATCTGGACAGCAATTTCTTTTGATCTTGAAT 1942  
Db 30 TTGGCGCTACTGCTCGTTGGGAATCTGGACAGCAATTTCTTTTGATCTTGAAT 89  
QY 1943 CGGATAACCATTTAGACAGCCCTCTAGTAAAGTATTGACAGATGCTCAACAGATGTTG 2002  
Db 90 CGGATAACCATTTAGACAGCCCTCTAGTAAAGTATTGACAGATGCTCAACAGATGTTG 149  
QY 2003 CTGCATGTTGGAAGAAATTCCTTTAGTACATACATTCCTGCTTCACTGATGACAA 2062  
Db 150 CTGCATGTTGGAAGAAATTCCTTTAGTACATACATTCCTGCTTCACTGATGACAA 209  
QY 2063 GAGGAAACGTTGTCATGATGATACCAAGAAAGGCAAGACACTTTAAACAGACTG 2122  
Db 210 GAGGAAACGTTGTCATGATGATACCAAGAAAGGCAAGACACTTTAAACAGACTG 269  
QY 2123 GGTTCCTCTCTCAG 2137  
Db 270 GGTTCCTCTCTCAG 284  
RESULT 5  
AAU01169 standard; cDNA; 591 BP.  
ID AAU01169;  
XX  
XX AAU01169;  
AC  
XX  
XX 21-NOV-2001 (first entry)  
DT  
XX  
XX Human reproductive system related antigen cDNA SEQ ID NO: 1170.  
DE  
XX  
XX Human, reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX PN WO200155320-A2.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01339.  
XX  
XX PR 31-JAN-2000; 2000US-0179065.  
XX  
XX PR 04-FEB-2000; 2000US-0180628.  
XX  
XX PR 24-FEB-2000; 2000US-0184664.  
XX  
XX PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0180076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 06-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241321.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251588.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI, 2001-465570/50.  
XX P-PSDB; AAM95199.  
XX  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen  
XX is used in preventing, treating or ameliorating a medical condition  
XX  
XX Claim 1, SEQ ID NO 1170; 1297bp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a





PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251989.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-48332/52.  
DR  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating testicular cancer -  
XX  
XX  
XX Claim 1; SEQ ID NO 298; 766bp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a cDNA of the  
CC invention.  
XX  
XX  
SQ Sequence 591 BP; 203 A; 100 C; 111 G; 176 T; 1 other;  
Query Match 11.7%; Score 255; DB 23; Length 591;  
Best Local Similarity 100.0%; Pred. No. 1,3e-115;  
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1863 TTGGCGTACTGCTGCTGTGGGAATCTGCGAGCAATTCCTTTTGTATCTTGAT 1942  
|||||

Db 30 TTGGCGTACTGCTGCTGTGGGAATCTGCGAGCAATTCCTTTTGTATCTTGAT 89  
QY 1943 CGATTAACCATTTAGCAGAGCTCTAGTAAAGATTGACAGATGCTCAACAGATGTT 2002  
Db 90 CGATTAACCATTTAGCAGAGCTCTAGTAAAGATTGACAGATGCTCAACAGATGTT 149  
QY 2003 CTGCATGTTGGAAGAAATTCCTTTAGTACATACATTCCTGCTTCAGTGTAGTACAA 2062  
Db 150 CTGCATGTTGGAAGAAATTCCTTTAGTACATACATTCCTGCTTCAGTGTAGTACAA 209  
QY 2063 GAGGAAACGTGTTTGCATCAGTTATACAGAAAGGCAAGAGACCTTTGAACACAGCTG 2122  
Db 210 GAGGAAACGTGTTTGCATCAGTTATACAGAAAGGCAAGAGACCTTTGAACACAGCTG 269  
QY 2123 GGTTCCTCTTCAC 2137  
Db 270 GGTTCCTCTTCAC 284  
RESULT 7  
AAL04793  
ID AAL04793 standard; DNA; 2194 BP.  
XX  
XX AAL04793;  
AC  
XX  
XX 21-NOV-2001 (first entry)  
DT  
XX  
XX Human reproductive system related antigen DNA SEQ ID NO: 7481.  
DE  
XX Human, reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.  
XX  
XX Homo sapiens.  
OS  
XX MO200155320-A2.  
PM  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01339.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0218290.  
PR 14-JUL-2000; 2000US-0218296.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.

```
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251868.
PR 05-DEC-2000; 2000US-0251869.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-46570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
XX Disclosure; SEQ ID NO 7481; 1297bp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
XX Sequence 2194 BP; 624 A; 349 C; 400 G; 821 T; 0 other;

Query Match 10.1%; Score 220; DB 22; Length 2194;
Best Local Similarity 100.0%; Pred. No. 2,6e-98;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1768 AAGTGGCCAGTTCTTGTCTACTTCACTAGTGCAGAGGGCGATATTGAAATGTG 1827
DB 1 AAGTGGCCAGTTCTTGTCTACTTCACTAGTGCAGAGGGCGATATTGAAATGTG 60
QY 1828 CAACATGTTATCAATTTTATCTTCTTACCATTTGATGATATGTAATGGG 1887
DB 61 CAACATGTTATCAATTTTATCTTCTTACCATTTGATGATATGTAATGGG 120
QY 1888 CGTACTGTCGTGTGTGGAACTGCGACAGCAATTTCTTTTATCTTGAATGGAT 1947
DB 121 CGTACTGTCGTGTGTGGAACTGCGACAGCAATTTCTTTTATCTTGAATGGAT 180
QY 1948 AACCATTTAGCAGAGCCTCTAGTAAAGATTGACAGATG 1987
DB 181 AACCATTTAGCAGAGCCTCTAGTAAAGATTGACAGATG 220

RESULT 8
ABL97687
```

ID ABL97687 standard; DNA; 2194 BP.  
XX ABL97687;  
AC  
XX  
XX  
DT 21-JUN-2002 (first entry)  
XX  
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2339.  
XX  
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ds.  
OS Homo sapiens.  
XX  
XX MO200155317-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01329.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225759.  
XX 14-AUG-2000; 2000US-0225759.  
XX 16-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 23-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 05-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 06-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234423.  
PR 21-SEP-2000; 2000US-0234474.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI, 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating testicular cancer  
XX  
XX Disclosure; SEQ ID NO 2339; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention.  
XX  
SQ Sequence 2194 BP; 624 A; 349 C; 400 G; 821 T; 0 other;  
  
Query Match 10.1%; Score 220; DB 23; Length 2194;  
Best Local Similarity 100.0%; Pred. No. 2,6e-98;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1768 AAGTGCACAGTTCTTGTGCTACTGAGTGCAGAGGCTGATTTGAAATGTG 1827  
Db 1 AAGTGCACAGTTCTTGTGCTACTGAGTGCAGAGGCTGATTTGAAATGTG 60  
  
QY 1828 CAACATGTTATCAATTTGATCTTCTTACCATTTGATGATTAATGATGGG 1887  
Db 61 CAACATGTTATCAATTTGATCTTCTTACCATTTGATGATTAATGATGGG 120  
  
QY 1888 CGTACTGCTGTGTGGGAACTACTGCGAGCAATTTCTTTTGTGATCTGGATGGAT 1947  
Db 121 CGTACTGCTGTGTGGGAACTACTGCGAGCAATTTCTTTTGTGATCTGGATGGAT 180  
  
QY 1948 AACCATTTAGCAGAGCTCTAGTAAAGTATGACAGATG 1987  
Db 181 AACCATTTAGCAGAGCTCTAGTAAAGTATGACAGATG 220  
  
RESULT 9  
AAL04794  
ID AAL04794 standard; DNA; 2195 BP.  
XX  
XX AAL04794;  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX Human reproductive system related antigen DNA SEQ ID NO: 7482.  
DE  
XX  
XX Human; reproductive system related antigen; reproductive system disorder;  
KM cancer; gene therapy; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W0200155320-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX

PF 17-JAN-2001; 2001MO-US01339.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 02-MAR-2000; 2000US-0184664.  
PR 16-MAR-2000; 2000US-0186350.  
PR 17-MAR-2000; 2000US-0198974.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227189.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249266.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen

PT is used in preventing, treating or ameliorating a medical condition  
XX  
XX Disclosure; SEQ ID NO 7482; 1297bp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.  
XX  
SQ Sequence 2195 BP; 624 A; 350 C; 400 G; 821 T; 0 other;  
Query Match 10.1%; Score 220; DB 22; Length 2195;  
Best Local Similarity 100.0%; Pred. No. 2.6e-98;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1768 AAGTGGCCAGTTTGTGTGCTTACTAGTAGCTGCGCAGAGGGGTGATATTGAAATGCG 1827  
DB 1 AAGTGGCCAGTTTGTGTGCTTACTAGTAGCTGCGCAGAGGGGTGATATTGAAATGCG 60  
QY 1828 CAACATGTTATCAATTTGATCTCTTACCATGATGATATGTCATCGAATTGGG 1887  
DB 61 CAACATGTTATCAATTTGATCTCTTACCATGATGATATGTCATCGAATTGGG 120  
QY 1888 CGTACTGTCGTCGTGTGGATAGTGGACAGACAAATTTCTTTTATCTTGATCGAT 1947  
DB 121 CGTACTGTCGTCGTGTGGATAGTGGACAGACAAATTTCTTTTATCTTGATCGAT 180  
QY 1948 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 1987  
DB 181 AACCATTTAGCACAGCCTCTAGTAAAGTATTGACAGATG 220  
RESULT 10  
ABL97688  
ID ABL97688 standard; DNA: 2195 BP.  
XX  
XX ABL97688;  
AC  
XX  
XX 21-JUN-2002 (first entry)  
DT  
XX  
XX Human testicular antigen encoding DNA fragment SEQ ID NO: 2340.  
DE  
XX  
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO200155317-A2.  
XX  
XX  
PD 02-AUG-2001.  
XX  
XX  
PF 17-JAN-2001; 2001WO-US01329.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 16-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.











```
PR 02-MAY-2001; 2001US-287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of
XX a genome, useful for detecting tissue-, pathology-, and
XX developmental-specific genes.
XX
XX Example 1; SEQ ID 15365; 47pp; English.
XX
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple
XX transcript units that populate a genome. The library comprises
XX several oligonucleotides, each capable of hybridizing selectively to a
XX set of messenger RNAs transcribed from a given transcription unit of
XX the genome, which encodes one or more messenger RNA splice variants.
XX The oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterizing the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialized mini
XX libraries to detect transcripts of a sub-transcriptome under a
XX particular biological or pathological state, and so allowing the
XX detection of tissue- and pathology-specific genes such as those genes
XX only expressed in specific tissue under a specific pathological
XX condition; to detect developmental specific genes; and to detect RNA
XX transcripts and splice variants of a transcriptome of a patient suffering
XX from a particular disorder. ABN27253 to ABN59589 represent
XX oligonucleotide sequences from rats, humans and mice, which are used in
XX the exemplification of the present invention.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WPIO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 60 BP; 19 A; 11 C; 18 G; 12 T; 0 other;
XX
XX Query Match 2.8%; Score 60; DB 24; Length 60;
XX Best Local Similarity 100.0%; Pred. No. 4.1e-19;
XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2065 GGAACGCTTTTCATGATGATACCAAGGAGGACCACTTTGAACACAGCTGGG 2124
XX DB 1 GGAACGCTTTTCATGATGATACCAAGGAGGACCACTTTGAACACAGCTGGG 60
XX
XX RESULT 14
XX AB199243
XX ID AB199243 standard; cDNA; 2187 BP.
XX
XX AC AB199243;
XX
XX DT 07-MAR-2002 (first entry)
XX
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:67.
XX
XX KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
XX OS Mus musculus.
XX
XX PN WO20018188-A2.
XX
XX PD 22-NOV-2001.
XX
XX PF 18-MAY-2001; 2001MO-JP04192.
XX
XX PR 18-MAY-2000; 2000JP-0145977.
XX
```

```
XX
XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
XX PI Ishikawa K, Asei S, Takahashi Y, Nagata T, Iehi Y;
XX WPI; 2002-034733/04.
XX
XX DR P-PSDB; ABB57045.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX expression levels of particular genes defined in the specification or
XX by determining the expression profile of a gene group comprising these
XX genes.
XX
XX Claim 2; Page 210-215; 2690pp; English.
XX
XX The present invention describes a method for examining ischemic
XX conditions, comprising measuring the expression levels of particular
XX genes (I) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (I). The method
XX is useful for examining the ischemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (AB199202 to AB199912, encoding
XX the protein sequences in ABB57020 to ABB57374) or by determining the
XX expression profile of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischemic condition-improving
XX drugs or therapeutics for ischemic diseases. AB199913 and AB199914
XX represent PCR primers for a mouse ischemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX
XX Sequence 2187 BP; 635 A; 409 C; 596 G; 547 T; 0 other;
XX
XX Query Match 1.1%; Score 23; DB 24; Length 2187;
XX Best Local Similarity 100.0%; Pred. No. 0.83;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 982 TTGATGGCTTGTGCTCAACAGG.1004
XX DB 681 TTGATGGCTTGTGCTCAACAGG 703
XX
XX RESULT 15
XX AAS53256
XX ID AAS53256 standard; DNA; 1842 BP.
XX
XX AC AAS53256;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Haemophilus influenzae DNA for cellular proliferation protein #38.
XX
XX KM Antisense; ds; prokaryotic cellular proliferation gene;
XX antibiotic; antibacterial; drug design.
XX
XX OS Haemophilus influenzae.
XX
XX PN WO200170955-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 21-MAR-2001; 2001MO-US09180.
XX
XX PR 21-MAR-2000; 2000US-191078P.
XX
XX PR 23-MAY-2000; 2000US-206848P.
XX
XX PR 26-MAY-2000; 2000US-207727P.
XX
XX PR 23-OCT-2000; 2000US-242578P.
XX
XX PR 27-NOV-2000; 2000US-253625P.
XX
XX PR 22-DEC-2000; 2000US-257931P.
XX
XX PR 16-FEB-2001; 2001US-269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 10:42:49 ; Search time 5458 Seconds  
(without alignments)  
11581.402 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 2172  
Sequence: 1 atgsgggatgagatctgga.....tagatgagatcatggaat 2172

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

## Database :

GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rod:\*

36: em\_htg\_mam:\*

37: em\_htg\_vit:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2172	100.0	2224	9 AY004154	AY004154 Homo sapi
2	1978	91.1	2411	9 AF262962	AF262962 Homo sapi
3	1719	79.1	2189	9 HSM802178	AI137462 Homo sapi
4	1265	58.2	1984	9 AK093439	AK093439 Homo sapi
5	273	12.6	139677	9 AC008914	AC008914 Homo sapi
6	273	12.6	176784	2 AC016632	AC016632 Homo sapi
7	273	12.6	182126	2 AC016639	AC016639 Homo sapi
8	273	12.6	204250	2 AC022265	AC022265 Homo sapi
9	174	8.0	204250	2 AC022265	AC022265 Homo sapi
10	58	2.7	524	4 AY100475	AY100475 Equus cab
11	53	2.4	3030	10 S75275	S75275 RVLG=vasa-1
12	50	2.3	1930	10 MUSDVH	D14859 Mouse mRNA
13	47	2.2	108795	2 AC124894	AC124894 Rattus no
14	47	2.2	114024	2 AC105884	AC105884 Rattus no
15	36	1.7	14350	1 AE013856	AE013856 Yersinia
16	36	1.7	335050	1 AJ414150	AJ414150 Yersinia
17	31	1.4	62682	2 AC109144	AC109144 Mus muscu
18	31	1.4	205057	2 AC108418	AC108418 Mus muscu
19	29	1.3	192541	2 AC127770	AC127770 Rattus no
20	28	1.3	2721	3 AB047382	AB047382 Hydra mag
21	27	1.2	108795	2 AC124894	AC124894 Rattus no
22	26	1.2	3731	10 MMU7376	AJ007376 Mus muscu
23	26	1.2	3767	10 BC021453	BC021453 Mus muscu
24	26	1.2	62682	2 AC109144	AC109144 Mus muscu
25	26	1.2	114024	2 AC105884	AC105884 Rattus no
26	25	1.2	178133	2 AC123925	AC123925 Mus muscu
27	23	1.1	1989	5 AB004836	AB004836 Gallus ga
28	23	1.1	2187	6 AX305316	AX305316 Sequence
29	23	1.1	2187	10 MUSRNAHLI	L25126 Mus musculu
30	31	3.132	23	3 AB047383	AB047383 Hydra mag
31	23	1.1	3185	10 MMDBRNAHL	Z38117 M.musculu
32	23	1.1	84767	9 AC004149	AC004149 Homo sapi
33	23	1.1	145870	2 AC068675	AC068675 Homo sapi
34	23	1.1	170530	2 AL833805	AL833805 Mus muscu
35	23	1.1	177744	2 AC073954	AC073954 Homo sapi
36	23	1.1	208955	2 AL808130	AL808130 Mus muscu
37	22	1.0	1512	5 AF479823	AF479823 Pantodon
38	22	1.0	43566	9 AF119709	AF119709 Homo sapi
39	22	1.0	43922	9 HS73E16	Z95330 Human DNA s
40	22	1.0	55237	2 AC119270	AC119270 Mus muscu
41	22	1.0	163640	9 AC093328	AC093328 Homo sapi
42	22	1.0	167794	2 AF215844	AF215844 Homo sapi
43	22	1.0	185319	2 AC110364	AC110364 Rattus no
44	22	1.0	237878	2 AC125202	AC125202 Mus muscu
45	21	1.0	1198	5 AF251800	AF251800 Dario dan
46	21	1.0	2492	5 DRY12007	Y12007 D.terio vlg
47	21	1.0	2865	5 AB005147	AB005147 Dario rer
48	21	1.0	10010	1 U32709	U32709 Haemophilus
49	21	1.0	70233	2 AC010831	AC010831 Homo sapi
50	21	1.0	84016	9 AL591375	AL591375 Human DNA
51	21	1.0	100288	2 AP001945	AP001945 Homo sapi
52	21	1.0	100853	9 AC094086	AC094086 Homo sapi
53	21	1.0	114724	2 AP004094	AP004094 Oryza sat
54	21	1.0	116773	2 AC009186	AC009186 Homo sapi
55	21	1.0	123593	9 AP001952	AP001952 Homo sapi
56	21	1.0	123661	9 AC096661	AC096661 Homo sapi
57	21	1.0	133695	2 AP003975	AP003975 Oryza sat
58	21	1.0	138159	2 AP002742	AP002742 Homo sapi
59	21	1.0	151019	9 AC026796	AC026796 Homo sapi
60	21	1.0	151663	2 AC102402	AC102402 Mus muscu
61	21	1.0	157814	2 AC016078	AC016078 Homo sapi
62	21	1.0	163793	2 AC109056	AC109056 Rattus no
63	21	1.0	167348	2 AP005104	AP005104 Oryza sat
64	21	1.0	172058	2 AC112815	AC112815 Rattus no
65	21	1.0	173316	2 AC110304	AC110304 Rattus no

66	21	1.0	174627	9	AC022138	AC022138 Homo sapi	139	20	0.9	161010	2	AC103631	AC103631 Mus muscu
C 67	21	1.0	179144	9	AC026746	AC026746 Homo sapi	140	20	0.9	162897	2	AC128474	AC128474 Rattus no
C 68	21	1.0	180303	2	AC018969	AC018969 Homo sapi	C 141	20	0.9	162955	2	AC092877	AC092877 Homo sapi
C 69	21	1.0	184826	2	AC122957	AC122957 Rattus no	C 142	20	0.9	163699	2	AC112462	AC112462 Rattus no
C 70	21	1.0	185825	10	AL671868	AL671868 Mouse DNA	C 143	20	0.9	163992	9	AC078889	AC078889 Homo sapi
C 71	21	1.0	196605	2	AC114368	AC114368 Rattus no	C 144	20	0.9	164229	9	AC006386	AC006386 Homo sapi
C 72	21	1.0	196716	2	AC026750	AC026750 Homo sapi	C 145	20	0.9	164808	9	AC025603	AC025603 Homo sapi
C 73	21	1.0	213351	2	AC026809	AC026809 Homo sapi	C 146	20	0.9	164744	2	AC025603	AC025603 Homo sapi
C 74	20	0.9	309	8	SVU76623	SVU76623 Sessantia ve	C 147	20	0.9	168081	2	AC083955	AC083955 Mus muscu
C 75	20	0.9	453	8	BSU78833	BSU78833 Bolusantus	C 148	20	0.9	169493	2	AL356605	AL356605 Homo DNA
C 76	20	0.9	630	8	CSU78838	CSU78838 Carmichaeli	C 149	20	0.9	169998	9	AC064846	AC064846 Homo sapi
C 77	20	0.9	714	11	PM12G5G	PM12G5G Mus muscullu	C 150	20	0.9	172919	9	AL807742	AL807742 Human DNA
C 78	20	0.9	737	10	MUSPRHA	MUSPRHA Mus muscullu	C 151	20	0.9	173377	2	AC114043	AC114043 Rattus no
C 79	20	0.9	1724	8	AF106843	AF106843 Peperomia	C 152	20	0.9	173852	2	AC095235	AC095235 Rattus no
C 80	20	0.9	2319	8	AY049285	AY049285 Arabidops	C 153	20	0.9	175601	2	AC103527	AC103527 Rattus no
C 81	20	0.9	2320	8	AY088132	AY088132 Arabidops	C 154	20	0.9	175998	2	AC119617	AC119617 Rattus no
C 82	20	0.9	2384	8	AB047381	AB047381 Hydra mag	C 155	20	0.9	177424	2	AC009832	AC009832 Homo sapi
C 83	20	0.9	2502	5	AF046043	AF046043 Xenopus l	C 156	20	0.9	178660	2	AC009852	AC009852 Homo sapi
C 84	20	0.9	2624	6	HSEMP42G1	HSEMP42G1 Homo sapi	C 157	20	0.9	180273	9	CNS01DU0	CNS01DU0 Homo chr
C 85	20	0.9	18946	3	CRC27D8	CRC27D8 Sequence	C 158	20	0.9	180797	2	AC116358	AC116358 Homo sapi
C 86	20	0.9	25201	2	AC092204	AC092204 Homo sapi	C 159	20	0.9	181716	2	AC113230	AC113230 Sus scrof
C 87	20	0.9	25347	3	CEC18D4	CEC18D4 Ceenorhadi	C 160	20	0.9	182660	2	AC012416	AC012416 Homo sapi
C 88	20	0.9	34841	3	CEC46F11	CEC46F11 Ceenorhadi	C 161	20	0.9	182972	2	AC115454	AC115454 Rattus no
C 89	20	0.9	34841	3	CEC46F11	CEC46F11 Ceenorhadi	C 162	20	0.9	183228	2	AC102671	AC102671 Mus muscu
C 90	20	0.9	34841	3	CEC46F11	CEC46F11 Ceenorhadi	C 163	20	0.9	183413	2	AC112000	AC112000 Homo sapi
C 91	20	0.9	56719	9	AC004058	AC004058 Homo sapi	C 164	20	0.9	183719	2	AC068842	AC068842 Homo sapi
C 92	20	0.9	56866	2	AC103759	AC103759 Homo sapi	C 165	20	0.9	183982	2	AC009825	AC009825 Homo sapi
C 93	20	0.9	59315	2	AC084169	AC084169 Homo sapi	C 166	20	0.9	184231	2	AC023651	AC023651 Homo sapi
C 94	20	0.9	61880	2	AC113129	AC113129 Mus muscu	C 167	20	0.9	184869	2	AC109695	AC109695 Rattus no
C 95	20	0.9	63910	10	AL646045	AL646045 Mouse DNA	C 168	20	0.9	185237	9	AC011124	AC011124 Homo sapi
C 96	20	0.9	64341	10	AL627323	AL627323 Mouse DNA	C 169	20	0.9	185286	10	AC006508	AC006508 Homo sapi
C 97	20	0.9	64371	2	AC100463	AC100463 Mus muscu	C 170	20	0.9	185755	9	AC105460	AC105460 Homo sapi
C 98	20	0.9	70540	2	AL136224	AL136224 Human DNA	C 171	20	0.9	185931	2	AL133554	AL133554 Homo sapi
C 99	20	0.9	74712	2	AC121278	AC121278 Mus muscu	C 172	20	0.9	186908	2	AC023770	AC023770 Homo sapi
C 100	20	0.9	89159	5	AL593845	AL593845 Zedraflish	C 173	20	0.9	187999	2	AC026928	AC026928 Homo sapi
C 101	20	0.9	90370	2	AC112325	AC112325 Rattus no	C 174	20	0.9	188107	2	AC113225	AC113225 Homo sapi
C 102	20	0.9	92346	9	AC074289	AC074289 Homo sapi	C 175	20	0.9	188116	10	AL732332	AL732332 Mouse DNA
C 103	20	0.9	93398	2	AP001863	AP001863 Homo sapi	C 176	20	0.9	189590	2	AC034284	AC034284 Homo sapi
C 104	20	0.9	94239	8	ATF14B22	ATF14B22 Arabidops	C 177	20	0.9	190727	9	AC068724	AC068724 Homo sapi
C 105	20	0.9	103517	8	AC073395	AC073395 Arabidops	C 178	20	0.9	194466	9	AL591046	AL591046 Human DNA
C 106	20	0.9	106108	2	AC107129	AC107129 Rattus no	C 179	20	0.9	195916	2	AC097079	AC097079 Rattus no
C 107	20	0.9	107148	2	AC117003	AC117003 Rattus no	C 180	20	0.9	199421	2	AC051636	AC051636 Homo sapi
C 108	20	0.9	107719	2	AC110847	AC110847 Rattus no	C 181	20	0.9	199654	2	AC010898	AC010898 Homo sapi
C 109	20	0.9	108127	2	AC125477	AC125477 Rattus no	C 182	20	0.9	200000	2	AC005289	AC005289 Homo sapi
C 110	20	0.9	108388	2	AP001821	AP001821 Homo sapi	C 183	20	0.9	202429	2	AC130160	AC130160 Rattus no
C 111	20	0.9	108451	2	AP002739	AP002739 Homo sapi	C 184	20	0.9	202490	2	AC121884	AC121884 Mus muscu
C 112	20	0.9	109082	2	AC026133	AC026133 Homo sapi	C 185	20	0.9	202849	2	AC130630	AC130630 Rattus no
C 113	20	0.9	110282	9	AC096563	AC096563 Homo sapi	C 186	20	0.9	207951	2	AC012264	AC012264 Homo sapi
C 114	20	0.9	113028	9	HS68D15	HS68D15 Human DNA	C 187	20	0.9	208035	2	AC006520	AC006520 Mus muscu
C 115	20	0.9	118067	2	AC114942	AC114942 Homo sapi	C 188	20	0.9	209392	2	AC113124	AC113124 Mus muscu
C 116	20	0.9	118777	10	AF163865	AF163865 Mus muscu	C 189	20	0.9	209817	9	CNS0180T	CNS0180T Human chr
C 117	20	0.9	119679	2	AC095884	AC095884 Rattus no	C 190	20	0.9	212560	2	AC103688	AC103688 Homo sapi
C 118	20	0.9	123905	2	AC116311	AC116311 Homo sapi	C 191	20	0.9	214441	2	AC122403	AC122403 Rattus no
C 119	20	0.9	125607	2	AP003913	AP003913 Oryza sat	C 192	20	0.9	216800	10	AL589744	AL589744 Mouse DNA
C 120	20	0.9	127701	2	AC091302	AC091302 Oryza sat	C 193	20	0.9	219062	2	AC102840	AC102840 Mus muscu
C 121	20	0.9	129584	2	AC101911	AC101911 Mus muscu	C 194	20	0.9	236723	2	AC068771	AC068771 Homo sapi
C 122	20	0.9	132179	2	AP002741	AP002741 Homo sapi	C 195	20	0.9	241664	2	AC124504	AC124504 Mus muscu
C 123	20	0.9	132179	2	AL672142	AL672142 Human DNA	C 196	20	0.9	250050	1	AL591976	AL591976 Listeria
C 124	20	0.9	133515	2	AL596091	AL596091 Homo sapi	C 197	20	0.9	251773	2	AC079129	AC079129 Mus muscu
C 125	20	0.9	134922	9	AL589678	AL589678 Human DNA	C 198	20	0.9	269616	2	AC124183	AC124183 Mus muscu
C 126	20	0.9	136257	9	AC068797	AC068797 Homo sapi	C 199	20	0.9	305346	2	AC125327	AC125327 Mus muscu
C 127	20	0.9	139373	9	AL596275	AL596275 Human DNA	C 200	20	0.9	319404	2	AC112130	AC112130 Homo sapi
C 128	20	0.9	143179	2	AC109346	AC109346 Homo sapi	C 201	20	0.9	325956	2	AL807385	AL807385 Mus muscu
C 129	20	0.9	144303	2	AC108758	AC108758 Oryza sat	C 202	20	0.9	363	9	CTU04446	CTU04446 Homo sapi
C 130	20	0.9	146315	2	AC103617	AC103617 Mus muscu	C 203	20	0.9	365	11	HS321W1	HS321W1 Homo sapi
C 131	20	0.9	152883	8	AC084319	AC084319 Oryza sat	C 204	20	0.9	489	8	AB076108S1	AB076108 Voivox ca
C 132	20	0.9	154286	9	AC016698	AC016698 Homo sapi	C 205	20	0.9	489	8	AB076110S1	AB076110 Voivox ca
C 133	20	0.9	154535	9	AC114283	AC114283 Homo sapi	C 206	20	0.9	599	3	CTI279788	CTI279788 Chrysosolm
C 134	20	0.9	155539	9	AL138721	AL138721 Human DNA	C 207	20	0.9	765	3	MSE6029	MSE6029 Manduca s
C 135	20	0.9	155779	2	AC068260	AC068260 Homo sapi	C 208	20	0.9	927	3	AF143849	AF143849 Adineta v
C 136	20	0.9	156956	2	AC010080	AC010080 Homo sapi	C 209	20	0.9	983	8	AB076115S1	AB076115 Gonium mu
C 137	20	0.9	157165	2	AC104112	AC104112 Homo sapi	C 210	20	0.9	1122	6	AX110914	AX110914 Sequence
C 138	20	0.9	160324	9	AC079117	AC079117 Homo sapi	C 211	20	0.9	1128	8	AB013999	AB013999 Voivox ca

212	19	0.9	1128	8	AB014006	AB014006 Pleodorina	285	19	0.9	13891	3	AC006665	AC006665 Caenorhab
213	19	0.9	1128	8	AB014007	AB014007 Eudorina	286	19	0.9	14087	1	AE006177	AE006177 Paeeteurei
214	19	0.9	1128	8	AB014008	AB014008 Eudorina	287	19	0.9	14560	10	AE006152	AE006152 Rattus no
215	19	0.9	1128	8	AB014009	AB014009 Eudorina	288	19	0.9	18705	3	CEY26E6A	CEY26E6A
216	19	0.9	1128	8	AB014010	AB014010 Eudorina	289	19	0.9	18909	10	MMU223837	MMU223837 Caenorhab
217	19	0.9	1128	8	AB014013	AB014013 Eudorina	290	19	0.9	21230	9	HS381G6	HS381G6 Mus muscu
218	19	0.9	1128	8	AB014029	AB014029 Volvulina	291	19	0.9	23703	10	AF220365	AF220365 Mus muscu
219	19	0.9	1128	8	AB014033	AB014033 Eudorina	292	19	0.9	27970	8	AP004518	AP004518 lotus jap
220	19	0.9	1128	8	AB044172	AB044172 Yamagishi	293	19	0.9	29987	3	SC8156	SC8156 Caenorhabd
221	19	0.9	1128	8	AB044533	AB044533 Lophomonas	294	19	0.9	31001	8	CEM07G4	CEM07G4
222	19	0.9	1128	8	AB047068	AB047068 Eudorina	295	19	0.9	32193	3	AC105348	AC105348
223	19	0.9	1128	8	AB047069	AB047069 Eudorina	296	19	0.9	33930	3	CEK04H4	CEK04H4
224	19	0.9	1128	8	AB047070	AB047070 Eudorina	297	19	0.9	35250	8	AY007366	AY007366
225	19	0.9	1128	8	AB047071	AB047071 Eudorina	298	19	0.9	36948	2	AC005499	AC005499 Drosophi
226	19	0.9	1128	8	AB047072	AB047072 Eudorina	299	19	0.9	38968	2	AC005793	AC005793 Homo sapi
227	19	0.9	1128	8	AB047073	AB047073 Eudorina	300	19	0.9	42074	9	AC004202	AC004202 Homo sapi
228	19	0.9	1128	8	AB076112	AB076112 Volvox gi	301	19	0.9	45977	2	AC0020130	AC0020130 Drosophi
229	19	0.9	1128	8	AB076113	AB076113 Volvox ob	302	19	0.9	46419	9	AC000058	AC000058 Homo sapi
230	19	0.9	1128	8	AB076114	AB076114 Volvox af	303	19	0.9	48768	9	AL157933	AL157933
231	19	0.9	1128	8	AB076120	AB076120 Vitreocchl	304	19	0.9	50000	6	AX393458	AX393458 Human DNA
232	19	0.9	1128	8	AB076121	AB076121 Vitreocchl	305	19	0.9	51417	2	AC119930	AC119930
233	19	0.9	1128	8	AB076122	AB076122 Vitreocchl	306	19	0.9	56835	2	AC025369	AC025369 Mus muscu
234	19	0.9	1132	10	BC023073	BC023073 Mus muscu	307	19	0.9	57339	2	AC044852	AC044852 Homo sapi
235	19	0.9	1290	6	AX109755	AX109755 Sequence	308	19	0.9	58484	9	AL137160	AL137160 Human DNA
236	19	0.9	1291	6	AX109756	AX109756 Sequence	309	19	0.9	58623	2	AC101064	AC101064
237	19	0.9	1337	9	AK024583	AK024583 Homo sapi	310	19	0.9	60310	2	AL391647	AL391647 Human DNA
238	19	0.9	1479	6	AX194078	AX194078 Sequence	311	19	0.9	60514	2	AC014121	AC014121 Drosophi
239	19	0.9	1586	9	AF520608	AF520608 Sparus au	312	19	0.9	60555	2	AC116907	AC116907 Homo sapi
240	19	0.9	1710	10	BC009141	BC009141 Mus muscu	313	19	0.9	63650	2	AC086342	AC086342
241	19	0.9	1752	8	AF096248	AF096248 Lycopersi	314	19	0.9	64923	2	AC002545	AC002545 Homo sapi
242	19	0.9	1898	9	AK001245	AK001245 Homo sapi	315	19	0.9	65386	2	AC1002532	AC1002532 Homo sapi
243	19	0.9	2019	6	AX489264	AX489264 Sequence	316	19	0.9	65972	2	AC124827	AC124827 Mus muscu
244	19	0.9	2123	9	AK093661	AK093661 Homo sapi	317	19	0.9	66065	2	AC090169	AC090169 Homo sapi
245	19	0.9	2128	9	AK027439	AK027439 Homo sapi	318	19	0.9	66353	2	AC095348	AC095348
246	19	0.9	2201	9	AF061337	AF061337 Homo sapi	319	19	0.9	67201	9	AL136169	AL136169
247	19	0.9	2293	9	AK027471	AK027471 Homo sapi	320	19	0.9	67246	2	AC105035	AC105035 Homo sapi
248	19	0.9	2306	9	AK027654	AK027654 Homo sapi	321	19	0.9	68970	2	AC090378	AC090378 Homo sapi
249	19	0.9	2319	9	HSAP000985	HSAP000985 Mus muscu	322	19	0.9	71468	2	AC102077	AC102077 Mus muscu
250	19	0.9	2388	10	BC030895	BC030895 Mus muscu	323	19	0.9	71548	2	AC097573	AC097573 Rattus no
251	19	0.9	2487	9	AK027425	AK027425 Homo sapi	324	19	0.9	71585	2	AC119959	AC119959 Mus muscu
252	19	0.9	2508	9	BC011819	BC011819 Homo sapi	325	19	0.9	73252	8	AC021199	AC021199 Arabidops
253	19	0.9	2556	9	AK091616	AK091616 Homo sapi	326	19	0.9	73259	2	AC084006	AC084006 Homo sapi
254	19	0.9	2722	6	AX327460	AX327460 Sequence	327	19	0.9	74686	2	AC104280	AC104280 Oryza sat
255	19	0.9	2838	10	AF159131	AF159131 Mus muscu	328	19	0.9	74989	2	AC023333	AC023333 Homo sapi
256	19	0.9	3041	9	AK027501	AK027501 Homo sapi	329	19	0.9	75605	2	AC111486	AC111486 Rattus no
257	19	0.9	3093	6	AX026741	AX026741 Sequence	330	19	0.9	77075	2	AC130514	AC130514 Rattus no
258	19	0.9	3123	9	AK001652	AK001652 Homo sapi	331	19	0.9	79526	9	AC025244	AC025244 Homo sapi
259	19	0.9	3158	9	HSU50553	HSU50553 Homo sapien	332	19	0.9	83059	9	AP001435	AP001435 Homo sapi
260	19	0.9	3408	9	HSAP000983	HSAP000983 Mus muscu	333	19	0.9	83550	9	AL390918	AL390918 Human DNA
261	19	0.9	3550	6	AX099520	AX099520 Sequence	334	19	0.9	83673	9	HSDA79C13	HSDA79C13
262	19	0.9	3634	9	AK096337	AK096337 Homo sapi	335	19	0.9	84277	2	AC105679	AC105679 Rattus no
263	19	0.9	3862	6	AX026746	AX026746 Sequence	336	19	0.9	86362	2	AC129445	AC129445
264	19	0.9	3862	6	AF279656	AF279656 Homo sapi	337	19	0.9	89609	2	AC106134	AC106134 Rattus no
265	19	0.9	4250	9	AB037789	AB037789 Homo sapi	338	19	0.9	90171	9	AC095040	AC095040 Homo sapi
266	19	0.9	4308	3	AV070840	AV070840 Drosophi	339	19	0.9	90623	9	AC090993	AC090993 Homo sapi
267	19	0.9	4416	9	HSAP000984	HSAP000984 Mus muscu	340	19	0.9	96363	9	AF230637	AF230637 Homo sapi
268	19	0.9	5197	8	LES6379	LES6379 Lycopersi	341	19	0.9	91045	8	AP003371	AP003371 Oryza sat
269	19	0.9	5322	9	HSAP000982	HSAP000982 Mus muscu	342	19	0.9	91640	2	AC123295	AC123295 Rattus no
270	19	0.9	6060	9	AK027867	AK027867 Homo sapi	343	19	0.9	95105	9	AC016561	AC016561 Homo sapi
271	19	0.9	6073	2	AC012949	AC012949 Drosophi	344	19	0.9	97687	9	AC008642	AC008642 Homo sapi
272	19	0.9	7130	6	AX175460	AX175460 Sequence	345	19	0.9	98398	9	AL1731533	AL1731533 Human DNA
273	19	0.9	7130	14	AF356698	AF356698 Porcine e	346	19	0.9	100000	9	AP000008	AP000008 Homo sapi
274	19	0.9	7161	9	AB011168	AB011168 Homo sapi	347	19	0.9	100000	9	AP0000149	AP0000149 Homo sapi
275	19	0.9	7174	9	HSMB804580	HSMB804580 Homo sapi	348	19	0.9	100000	2	AP000518	AP000518 Homo sapi
276	19	0.9	7238	3	CECL82G	CECL82G C. elegans	349	19	0.9	100516	2	HS171M_3	HS171M_3
277	19	0.9	9653	2	AC020842	AC020842 Mus muscu	350	19	0.9	101539	8	ATTSPI_19	ATTSPI_19
278	19	0.9	10029	1	AE010558	AE010558 Rubobacte	351	19	0.9	101624	2	AC102968	AC102968
279	19	0.9	10252	1	AE000075	AE000075 Ruzobium	352	19	0.9	102370	9	HSJ976013	HSJ976013 Human DNA
280	19	0.9	10477	1	AE007510	AE007510 Streptococ	353	19	0.9	103332	2	AC0933793	AC0933793 Homo sapi
281	19	0.9	10478	1	AE008567	AE008567 Streptococ	354	19	0.9	103473	2	AC126952	AC126952 Rattus no
282	19	0.9	10753	6	AX111980	AX111980 Sequence	355	19	0.9	104311	2	AC105867	AC105867 Rattus no
283	19	0.9	10753	6	AX175459	AX175459 Sequence	356	19	0.9	106552	2	AL805913	AL805913 Homo sapi
284	19	0.9	12813	1	AE009776	AE009776 Pyrobaculi	357	19	0.9	106582	9	AL360215	AL360215 Human DNA

[illegible]



C 504	19	0.9 16384	9	AC009180	AC009180 Homo sapi	C 577	19	0.9 177835	3	AC091127	AC091127 Drosophila
C 505	19	0.9 163842	2	AC096224	AC096224 Rattus no	C 578	19	0.9 178257	2	AC011781	AC011781 Homo sapi
C 506	19	0.9 163881	2	AC007436	AC007436 Homo sapi	C 579	19	0.9 178719	2	AC114135	AC114135 Rattus no
C 507	19	0.9 163978	2	AC126066	AC126066 Rattus no	C 580	19	0.9 178965	2	AC096167	AC096167 Rattus no
C 508	19	0.9 164343	2	AC024593	AC024593 Homo sapi	C 581	19	0.9 178986	2	AC046200	AC046200 Homo sapi
C 509	19	0.9 164743	2	AP005636	AP005636 Oryza sat	C 582	19	0.9 179387	2	AC113689	AC113689 Rattus no
C 510	19	0.9 164773	2	AC091543	AC091543 Felis cat	C 583	19	0.9 179647	2	AC108124	AC108124 Homo sapi
C 511	19	0.9 164810	2	AC109166	AC109166 Mus muscu	C 584	19	0.9 179856	3	AC010067	AC010067 Drosophila
C 512	19	0.9 165024	2	AC093707	AC093707 Pan trogl	C 585	19	0.9 180293	2	AC108143	AC108143 Homo sapi
C 513	19	0.9 165093	2	AC069444	AC069444 Homo sapi	C 586	19	0.9 180304	2	AC012541	AC012541 Homo sapi
C 514	19	0.9 165385	2	AC103649	AC103649 Mus muscu	C 587	19	0.9 180773	9	AC092969	AC092969 Homo sapi
C 515	19	0.9 165639	2	AC099623	AC099623 Mus muscu	C 588	19	0.9 181245	9	AC093881	AC093881 Homo sapi
C 516	19	0.9 166421	8	AP003106	AP003106 Oryza sat	C 589	19	0.9 181512	9	AC116017	AC116017 Homo sapi
C 517	19	0.9 166960	9	AC117513	AC117513 Homo sapi	C 590	19	0.9 181786	2	AL845506	AL845506 Mus muscu
C 518	19	0.9 167132	9	AC069154	AC069154 Homo sapi	C 591	19	0.9 181827	2	AC055854	AC055854 Homo sapi
C 519	19	0.9 167234	9	AC009729	AC009729 Homo sapi	C 592	19	0.9 183054	2	AC116074	AC116074 Rattus no
C 520	19	0.9 167275	9	AC008782	AC008782 Homo sapi	C 593	19	0.9 183084	2	AC131222	AC131222 Rattus no
C 521	19	0.9 167658	2	AL772326	AL772326 Mus muscu	C 594	19	0.9 183244	2	AC027188	AC027188 Homo sapi
C 522	19	0.9 167705	2	AC074252	AC074252 Homo sapi	C 595	19	0.9 183451	9	AC005972	AC005972 Homo sapi
C 523	19	0.9 167945	2	AC129291	AC129291 Mus muscu	C 596	19	0.9 183547	2	AC009551	AC009551 Homo sapi
C 524	19	0.9 168105	9	AC093815	AC093815 Homo sapi	C 597	19	0.9 183659	9	AC068472	AC068472 Homo sapi
C 525	19	0.9 168413	2	AC009877	AC009877 Homo sapi	C 598	19	0.9 183698	10	AL606464	AL606464 Mouse DNA
C 526	19	0.9 168437	2	AC124948	AC124948 Homo sapi	C 599	19	0.9 183838	9	AC104332	AC104332 Homo sapi
C 527	19	0.9 168505	9	AC104457	AC104457 Homo sapi	C 600	19	0.9 184024	2	AC127121	AC127121 Rattus no
C 528	19	0.9 168545	9	HS287820	HS287820 Homo sapi	C 601	19	0.9 184172	2	AC114244	AC114244 Rattus no
C 529	19	0.9 168954	2	AC099706	AC099706 Mus muscu	C 602	19	0.9 184268	2	AC011401	AC011401 Homo sapi
C 530	19	0.9 169192	2	AC114672	AC114672 Mus muscu	C 603	19	0.9 184285	9	AC007365	AC007365 Homo sapi
C 531	19	0.9 169237	2	AP004869	AP004869 Oryza sat	C 604	19	0.9 184391	9	AL592043	AL592043 Human DNA
C 532	19	0.9 169354	2	AC099598	AC099598 Mus muscu	C 605	19	0.9 184427	14	EHV20824	EHV20824 Equine hepr
C 533	19	0.9 169816	2	AC118600	AC118600 Mus muscu	C 606	19	0.9 184579	2	AL824708	AL824708 Dario rex
C 534	19	0.9 169841	9	AL607043	AL607043 Human DNA	C 607	19	0.9 184841	9	AC016902	AC016902 Homo sapi
C 535	19	0.9 170058	2	AC124905	AC124905 Equus cab	C 608	19	0.9 184869	2	AC015622	AC015622 Homo sapi
C 536	19	0.9 170133	2	AC107335	AC107335 Rattus no	C 609	19	0.9 185051	2	AC127098	AC127098 Rattus no
C 537	19	0.9 170393	2	AC026473	AC026473 Homo sapi	C 610	19	0.9 185121	2	AC069105	AC069105 Homo sapi
C 538	19	0.9 170458	2	AC034127	AC034127 Homo sapi	C 611	19	0.9 185507	2	AC107562	AC107562 Rattus no
C 539	19	0.9 170795	2	AC021346	AC021346 Homo sapi	C 612	19	0.9 185528	2	AC113643	AC113643 Rattus no
C 540	19	0.9 170891	2	AC002118	AC002118 Homo sapi	C 613	19	0.9 185569	2	AP003099	AP003099 Homo sapi
C 541	19	0.9 170962	2	AL389927	AL389927 Homo sapi	C 614	19	0.9 185666	2	AC048389	AC048389 Homo sapi
C 542	19	0.9 170970	2	AL512510	AL512510 Human DNA	C 615	19	0.9 185709	2	AC094466	AC094466 Rattus no
C 543	19	0.9 171316	2	AL591712	AL591712 Mus muscu	C 616	19	0.9 185821	2	AL844866	AL844866 Mus muscu
C 544	19	0.9 171400	2	AC129698	AC129698 Rattus no	C 617	19	0.9 186049	2	AC117695	AC117695 Mus muscu
C 545	19	0.9 171502	2	AC068638	AC068638 Homo sapi	C 618	19	0.9 186709	2	AC113713	AC113713 Rattus no
C 546	19	0.9 171589	2	AC073567	AC073567 Homo sapi	C 619	19	0.9 187332	2	AC101292	AC101292 Mus muscu
C 547	19	0.9 171613	2	AC109435	AC109435 Rattus no	C 620	19	0.9 187359	2	AC068866	AC068866 Homo sapi
C 548	19	0.9 172266	3	AC092190	AC092190 Drosophila	C 621	19	0.9 187509	2	AC114370	AC114370 Rattus no
C 549	19	0.9 172360	2	AC024552	AC024552 Homo sapi	C 622	19	0.9 187680	2	AC126244	AC126244 Mus muscu
C 550	19	0.9 172724	2	AC109968	AC109968 Rattus no	C 623	19	0.9 187908	9	HS495010	HS495010 Human DNA
C 551	19	0.9 172931	9	AC006581	AC006581 Homo sapi	C 624	19	0.9 187942	9	AC104639	AC104639 Homo sapi
C 552	19	0.9 172965	2	AC023367	AC023367 Homo sapi	C 625	19	0.9 187976	2	AC127750	AC127750 Rattus no
C 553	19	0.9 172993	2	CNS01RM9	AL161665 Human chr	C 626	19	0.9 188018	2	AC021525	AC021525 Homo sapi
C 554	19	0.9 173113	2	AC092529	AC092529 Papio cyn	C 627	19	0.9 188207	2	AC010233	AC010233 Homo sapi
C 555	19	0.9 173133	2	AC016166	AC016166 Homo sapi	C 628	19	0.9 188425	2	AC026554	AC026554 Homo sapi
C 556	19	0.9 173201	2	AC008169	AC008169 Homo sapi	C 629	19	0.9 188697	2	AC125577	AC125577 Rattus no
C 557	19	0.9 173289	2	AC096462	AC096462 Rattus no	C 630	19	0.9 188871	10	AL713870	AL713870 Mouse DNA
C 558	19	0.9 173348	2	AC115499	AC115499 Rattus no	C 631	19	0.9 189056	2	AC024519	AC024519 Homo sapi
C 559	19	0.9 173450	9	AC091769	AC091769 Homo sapi	C 632	19	0.9 189092	2	AC011759	AC011759 Drosophila
C 560	19	0.9 173507	9	AC006385	AC006385 Homo sapi	C 633	19	0.9 189355	2	AP000921	AP000921 Homo sapi
C 561	19	0.9 173597	2	AC025749	AC025749 Homo sapi	C 634	19	0.9 189577	2	AC091199	AC091199 Homo sapi
C 562	19	0.9 173635	2	AC123432	AC123432 Rattus no	C 635	19	0.9 189829	2	AC113069	AC113069 Mus muscu
C 563	19	0.9 173677	9	AC084768	AC084768 Homo sapi	C 636	19	0.9 190338	2	AC130145	AC130145 Rattus no
C 564	19	0.9 174393	2	AC011720	AC011720 Homo sapi	C 637	19	0.9 190626	2	AC062036	AC062036 Homo sapi
C 565	19	0.9 174396	2	AC073076	AC073076 Homo sapi	C 638	19	0.9 191058	2	AC105910	AC105910 Homo sapi
C 566	19	0.9 175131	2	AC114820	AC114820 Mus muscu	C 639	19	0.9 191704	2	AC053529	AC053529 Homo sapi
C 567	19	0.9 175210	9	CNS07EG5	AL590328 Human chr	C 640	19	0.9 191728	2	AL512265	AL512265 Homo sapi
C 568	19	0.9 175372	2	AC009659	AC009659 Homo sapi	C 641	19	0.9 191923	2	AP001846	AP001846 Homo sapi
C 569	19	0.9 175466	2	AC016419	AC016419 Homo sapi	C 642	19	0.9 192399	3	AC010032	AC010032 Drosophila
C 570	19	0.9 175594	9	AL162503	AL162503 Human DNA	C 643	19	0.9 192748	2	AC113326	AC113326 Mus muscu
C 571	19	0.9 175647	2	AC105414	AC105414 Homo sapi	C 644	19	0.9 192869	2	AC102767	AC102767 Mus muscu
C 572	19	0.9 175878	2	AC130047	AC130047 Rattus no	C 645	19	0.9 192967	2	AC021992	AC021992 Homo sapi
C 573	19	0.9 175951	9	AC011998	AC011998 Homo sapi	C 646	19	0.9 193302	2	AC128553	AC128553 Mus muscu
C 574	19	0.9 176104	9	AC007163	AC007163 Homo sapi	C 647	19	0.9 193453	2	AC022273	AC022273 Homo sapi
C 575	19	0.9 176493	9	AC016682	AC016682 Homo sapi	C 648	19	0.9 193761	9	AL445183	AL445183 Human DNA
C 576	19	0.9 177102	9	HSJ1069C8	AL078623 Human DNA	C 649	19	0.9 193774	9	AL445183	AL445183 Human DNA

650	19	0.9	193872	9	AC087369	AC087369 Homo sapi	c 723	19	0.9	238845	2	AC103330	AC103330 Rattus no
c 651	19	0.9	193942	2	AC112925	AC112925 Mus muscu	c 724	19	0.9	247876	2	AL844873	AL844873 Mus muscu
c 652	19	0.9	195002	2	AC123812	AC123812 Mus muscu	c 725	19	0.9	248381	2	AC009937	AC009937 Homo sapi
c 653	19	0.9	195296	2	AC122234	AC122234 Mus muscu	c 726	19	0.9	249616	10	AE014183	AE014183 Mus muscu
c 654	19	0.9	195296	2	AC122234	AC122234 Mus muscu	c 727	19	0.9	259474	9	HUAC004605	HUAC004605 Homo sapi
c 655	19	0.9	195310	2	AC026606	AC026606 Homo sapi	c 728	19	0.9	263692	2	AC113251	AC113251 Rattus no
c 656	19	0.9	195326	2	AC024189	AC024189 Homo sapi	c 729	19	0.9	265786	9	AE006466	AE006466 Homo sapi
c 657	19	0.9	195682	9	AC120498	AC120498 Homo sapi	c 730	19	0.9	269616	2	AC124183	AC124183 Mus muscu
c 658	19	0.9	196363	9	AC096759	AC096759 Homo sapi	c 731	19	0.9	270100	2	AC126428	AC126428 Mus muscu
c 659	19	0.9	196472	2	AC011862	AC011862 Homo sapi	c 732	19	0.9	274388	3	AE003574	AE003574 Drosophi
c 660	19	0.9	196366	2	AC120676	AC120676 Rattus no	c 733	19	0.9	287178	2	AC125377	AC125377 Mus muscu
c 661	19	0.9	197351	2	AC127204	AC127204 Rattus no	c 734	19	0.9	291050	1	AE000982	AE000982 Drosophi
c 662	19	0.9	197403	2	AC123823	AC123823 Mus muscu	c 735	19	0.9	295225	3	AE003461	AE003461 Homo sapi
c 663	19	0.9	198259	2	AC117648	AC117648 Mus muscu	c 736	19	0.9	301711	3	AE003470	AE003470 Drosophi
c 664	19	0.9	198582	9	AC005291	AC005291 Homo sapi	c 737	19	0.9	302300	1	AE003186	AE003186 Drosotridi
c 665	19	0.9	198958	9	AC020688	AC020688 Homo sapi	c 738	19	0.9	309883	3	AE003479	AE003479 Homo sapi
c 666	19	0.9	199016	2	AC021387	AC021387 Homo sapi	c 739	19	0.9	317463	2	AC091254	AC091254 Mus muscu
c 667	19	0.9	200339	10	AE003100	AE003100 Homo sapi	c 740	19	0.9	320040	6	A79350	A79350 Sequence 1
c 668	19	0.9	200703	2	AC122833	AC122833 Mus muscu	c 741	19	0.9	320040	6	A93002	A93002 Sequence 1
c 669	19	0.9	200843	2	AC123482	AC123482 Rattus no	c 742	19	0.9	320436	2	AL772140	AL772140 Homo sapi
c 670	19	0.9	200853	9	AC111119	AC111119 Mus muscu	c 743	19	0.9	325808	2	AC121787	AC121787 Mus muscu
c 671	19	0.9	201136	9	CNS01D50	AL121839 Human chr	c 744	19	0.9	340000	9	AP001684	AP001684 Homo sapi
c 672	19	0.9	201312	2	AP000959	AP000959 Homo sapi	c 745	19	0.9	340000	9	AP001707	AP001707 Homo sapi
c 673	19	0.9	201676	2	AC067900	AC067900 Homo sapi	c 746	19	0.9	340000	9	AP001727	AP001727 Homo sapi
c 674	19	0.9	202565	9	AL354696	AL354696 Sus scrof	c 747	19	0.9	340000	9	HS21C102	HS21C102 Homo sapi
c 675	19	0.9	205339	10	AC091484	AC091484 Homo sapi	c 748	19	0.9	349980	6	AX344565	AX344565 Sequence
c 676	19	0.9	205993	2	AC011952	AC011952 Homo sapi	c 749	19	0.9	349980	6	AX344566	AX344566 Sequence
c 677	19	0.9	207018	2	AC116539	AC116539 Drosophi	c 750	18	0.8	107	11	HS50D11R	HS50D11R Sequence
c 678	19	0.9	207162	2	AL831752	AL831752 Mus muscu	c 751	18	0.8	163	11	HUMSWX1698	HUMSWX1698 Human chrom
c 679	19	0.9	207268	2	AL671854	AL671854 Mus muscu	c 752	18	0.8	174	6	AX088553	AX088553 Sequence
c 680	19	0.9	207534	2	AC073854	AC073854 Homo sapi	c 753	18	0.8	250	14	AF106188	AF106188 HIV-1 str
c 681	19	0.9	207534	2	AC024698	AC024698 Homo sapi	c 754	18	0.8	256	9	HS50B8R	HS50B8R HIV-1 str
c 682	19	0.9	207991	9	AC007790	AC007790 Homo sapi	c 755	18	0.8	339	6	AR203677	AR203677 Sequence
c 683	19	0.9	208265	10	AL606907	AL606907 Homo sapi	c 756	18	0.8	340	3	AB047387	AB047387 Dugesia d
c 684	19	0.9	208398	2	AC123245	AC123245 Rattus no	c 757	18	0.8	390	14	AB081225	AB081225 Human imm
c 685	19	0.9	208936	10	AL671478	AL671478 Mus muscu	c 758	18	0.8	390	14	AF015397	AF015397 HIV-1 str
c 686	19	0.9	208936	2	AC113181	AC113181 Mus muscu	c 759	18	0.8	390	14	AF015398	AF015398 HIV-1 str
c 687	19	0.9	209910	2	AC105301	AC105301 Mus muscu	c 760	18	0.8	390	14	AF015399	AF015399 HIV-1 str
c 688	19	0.9	210813	10	AL627070	AL627070 Mouse DNA	c 761	18	0.8	390	14	AF062042	AF062042 HIV-1 iso
c 689	19	0.9	210933	2	AC027394	AC027394 Homo sapi	c 762	18	0.8	390	14	AF062043	AF062043 HIV-1 iso
c 690	19	0.9	211337	2	AC021136	AC021136 Homo sapi	c 763	18	0.8	390	14	AF062044	AF062044 HIV-1 iso
c 691	19	0.9	211433	2	AL805896	AL805896 Mus muscu	c 764	18	0.8	392	11	G21206	G21206 human STR W
c 692	19	0.9	211844	2	AC048370	AC048370 Homo sapi	c 765	18	0.8	396	14	HIVISIMWN	HIVISIMWN human immun
c 693	19	0.9	212529	2	AC115948	AC115948 Mus muscu	c 766	18	0.8	421	6	AX302796	AX302796 Sequence
c 694	19	0.9	212560	2	AC103688	AC103688 Homo sapi	c 767	18	0.8	421	11	G07123	G07123 human STR W
c 695	19	0.9	213477	2	AC098347	AC098347 Rattus no	c 768	18	0.8	438	14	AF481781	AF481781 HIV-1 iso
c 696	19	0.9	213943	2	AC015465	AC015465 Homo sapi	c 769	18	0.8	452	6	AX110906	AX110906 Sequence
c 697	19	0.9	214572	10	AL450341	AL450341 Mouse DNA	c 770	18	0.8	532	6	AR138292	AR138292 Sequence
c 698	19	0.9	214596	2	AC122233	AC122233 Mus muscu	c 771	18	0.8	579	11	G66094	G66094 sy1028 Misc
c 699	19	0.9	216457	9	CNS01RGI	AL157689 Human chr	c 772	18	0.8	599	6	AX331374	AX331374 Sequence
c 700	19	0.9	217586	2	AC016076	AC016076 Homo sapi	c 773	18	0.8	710	9	HSX316229	HSX316229 Homo sapi
c 701	19	0.9	218565	3	AC009257	AC009257 Drosophi	c 774	18	0.8	717	11	PM4H3G	PM4H3G Penicilli
c 702	19	0.9	218956	9	CNS00008	AL049831 Human chr	c 775	18	0.8	721	14	AF051470	AF051470 HIV-1 iso
c 703	19	0.9	219145	2	AC102087	AC102087 Mus muscu	c 776	18	0.8	732	14	AF196674	AF196674 HIV-1 iso
c 704	19	0.9	219639	2	AC105206	AC105206 Homo sapi	c 777	18	0.8	836	9	HUMIGHAT	HUMIGHAT Human Ig ab
c 705	19	0.9	219825	10	AC098886	AC098886 Mus muscu	c 778	18	0.8	882	11	CNS06EXZ	AL395709 T7 end of
c 706	19	0.9	220407	2	AC112792	AC112792 Mouse DNA	c 779	18	0.8	895	3	AF027724	AF027724 Sequence
c 707	19	0.9	220450	2	HS53110	AL127892 Mus muscu	c 780	18	0.8	909	6	AX463545	AX463545 Sequence
c 708	19	0.9	220568	9	AC108048	AC108048 Homo sapi	c 781	18	0.8	912	8	AF404858	AF404858 Arbidops
c 709	19	0.9	220568	10	AL604045	AL604045 Mouse DNA	c 782	18	0.8	957	6	AX241558	AX241558 Sequence
c 710	19	0.9	224360	2	AC090670	AC090670 Homo sapi	c 783	18	0.8	1004	11	CNS06FTM	AL396056 T7 end of
c 711	19	0.9	224613	2	AC126555	AC126555 Mus muscu	c 784	18	0.8	1018	6	AX109947	AX109947 Sequence
c 712	19	0.9	226305	3	AE003773	AE003773 Drosophi	c 785	18	0.8	1028	11	CNS06IMR	AL000489 T7 end of
c 713	19	0.9	226905	2	AC087140	AC087140 Mus muscu	c 786	18	0.8	1105	3	AY119531	AY119531 Gallus
c 714	19	0.9	229277	2	HSUS2111	U52111 Homo sapien	c 787	18	0.8	1125	5	AB0110995	AB0110995 Gallus ga
c 715	19	0.9	232807	9	SPNEU1901	AL49923 Streptoco	c 788	18	0.8	1128	8	AB014000	AB014000 Volvox di
c 716	19	0.9	234053	2	AC079427	AC079427 Mus muscu	c 789	18	0.8	1128	8	AB014019	AB014019 Goniium qu
c 717	19	0.9	235183	2	AC079427	AC079427 Mus muscu	c 790	18	0.8	1132	8	AB014020	AB014020 Goniium mu
c 718	19	0.9	235183	2	AC079427	AC079427 Mus muscu	c 791	18	0.8	1132	8	AF308585	AF308585 Trilicho
c 719	19	0.9	235733	2	AL356746	AL356746 Homo sapi	c 792	18	0.8	1150	3	AY088401	AY088401 Arabidops
c 720	19	0.9	236809	2	AC022098	AC022098 Homo sapi	c 793	18	0.8	1195	3	AF529422	AF529422 Bombyx mo
c 721	19	0.9	237931	9	AC124513	AC124513 Mus muscu	c 794	18	0.8	1222	8	CCY419134	CCY419134 Chorizand
c 722	19	0.9	238413	2	AC124513	AC124513 Mus muscu	c 795	18	0.8	1263	10	AF238223	AF238223 Rattus no

796	18	0.8	1277	6	AX109728	Sequence	869	18	0.8	1972	5	AF479820	AF479820 Cyprius
797	18	0.8	1331	8	AT056519	Encephala	870	18	0.8	2003	10	RNSMAD4	AF156730 Rattus no
798	18	0.8	1337	8	AF209597	Heteropyxa	871	18	0.8	2005	6	AR030509	AR030509 Sequence
799	18	0.8	1346	8	MSA248322	Medicago	872	18	0.8	2085	1	AY099293	AY099293 Bartonella
800	18	0.8	1348	8	PL1417571	Paradijsia	873	18	0.8	2105	1	STU94901	STU94901 Salmonella
801	18	0.8	1352	8	SSP419158	Streptococ	874	18	0.8	2130	10	MMU01841	MMU01841 Mus musculus
802	18	0.8	1369	8	PPR235477	Streptococ	875	18	0.8	2141	8	AF105110	AF105110 Nymphaea
803	18	0.8	1374	8	AF209564	Clarkia p	876	18	0.8	2196	6	AX088888	AX088888 Sequence
804	18	0.8	1380	8	ORU93825	U93825 Ophioglossu	877	18	0.8	2219	9	AB070050	AB070050 Macaca fa
805	18	0.8	1410	1	SAU309184	SAU309184 Staphyloc	878	18	0.8	2261	6	AX088887	AX088887 Sequence
806	18	0.8	1418	8	AF209621	Lythrum s	879	18	0.8	2317	3	AF004915	AF004915 Anopheles
807	18	0.8	1423	10	AF250142	Rattus no	880	18	0.8	2428	3	AGA010195	AGA010195 Anopheles
808	18	0.8	1426	8	AF209582	Epiobium	881	18	0.8	2480	5	AF134850	AF134850 Homo sapi
809	18	0.8	1428	8	AF484684	Ana p	882	18	0.8	2500	9	HUMCGMP	HUMCGMP Homo sapi
810	18	0.8	1428	10	AF317514	Cavia por	883	18	0.8	2501	9	HSSMAD7604	HSSMAD7604 Homo sapi
811	18	0.8	1429	8	MMU10374	Mus musculu	884	18	0.8	2548	8	AY102111	AY102111 Arabidops
812	18	0.8	1429	8	AF209637	Neurada p	885	18	0.8	2632	14	AB078707	AB078707 Human imm
813	18	0.8	1432	8	AF197601	Trigloch	886	18	0.8	2715	6	AX399012	AX399012 Sequence
814	18	0.8	1433	8	ATH010458	Arabidops	887	18	0.8	2715	8	PTCCPRGOL	PTCCPRGOL Arabidops
815	18	0.8	1438	8	PPR235574	Punica pr	888	18	0.8	2812	5	XLU92030	XLU92030 Xenopus lae
816	18	0.8	1439	8	AF209633	Moringa o	889	18	0.8	2857	9	S42457	S42457 CNGC=rod ph
817	18	0.8	1443	8	AF209686	Terminali	890	18	0.8	2871	9	HUMBZM02	HUMBZM02 Homo sapien
818	18	0.8	1443	8	AF209686	Terminali	891	18	0.8	2902	10	AF214732	AF214732 Mus muscu
819	18	0.8	1445	8	PPA419149	Pharus pa	892	18	0.8	2910	3	AY075460	AY075460 Arabidops
820	18	0.8	1453	8	EMO419136	Ecdicol	893	18	0.8	2955	10	AF464181	AF464181 Mus muscu
821	18	0.8	1453	8	NPAU3069	Neurada p	894	18	0.8	3029	9	HSU88879	HSU88879 Homo sapien
822	18	0.8	1453	8	QIN235576	Quisquali	895	18	0.8	3103	6	ARI03433	ARI03433 Sequence
823	18	0.8	1456	10	AF020039	Mus muscu	896	18	0.8	3103	6	AR201459	AR201459 Sequence
824	18	0.8	1464	8	AF197610	Oxortium	897	18	0.8	3111	6	ARI21107	ARI21107 Sequence
825	18	0.8	1467	8	AAA03072	Abroma an	898	18	0.8	3111	6	AX334395	AX334395 Sequence
826	18	0.8	1467	8	NNE235535	Metroside	899	18	0.8	3111	9	AF010193	AF010193 Homo sapi
827	18	0.8	1471	8	AF197600	Petromogel	900	18	0.8	3157	8	S51252	S51252 Neurospora
828	18	0.8	1474	8	AF209677	Smilax gl	901	18	0.8	3165	3	AY089433	AY089433 Arabidops
829	18	0.8	1474	8	SWA235606	Spachibhy	902	18	0.8	3222	8	ATT83178	ATT83178 Arabidops
830	18	0.8	1475	8	JGA419143	Jolnville	903	18	0.8	3321	6	BD004082	BD004082 Apoptosis
831	18	0.8	1480	8	CEN419133	Chorizand	904	18	0.8	3513	6	AX277590	AX277590 Sequence
832	18	0.8	1480	8	HNE419132	Hypolytru	905	18	0.8	3513	9	HSMB00944	HSMB00944 Homo sapi
833	18	0.8	1491	8	AF168906	Cyperus a	906	18	0.8	3522	5	XELSTRGE	XELSTRGE Xenopus lae
834	18	0.8	1491	8	AF168915	Gymnstrac	907	18	0.8	3681	10	MSMAD7	MSMAD7 Mus muscu
835	18	0.8	1495	14	HIVG109	L11765 Human immu	908	18	0.8	4046	9	HSMB00935	HSMB00935 Homo sapi
836	18	0.8	1497	8	AAFA17568	Agapannu	909	18	0.8	4311	10	AF159626	AF159626 Rattus no
837	18	0.8	1500	8	AF168887	Apogoneto	910	18	0.8	4477	9	AK024475	AK024475 Homo sapi
838	18	0.8	1500	8	AF168894	Chlorophy	911	18	0.8	4528	10	AY013805	AY013805 Mus muscu
839	18	0.8	1500	8	AF168896	Colocasia	912	18	0.8	5001	8	CHPM5KB	AJ222802 Proteothec
840	18	0.8	1500	8	AF168898	Costus ba	913	18	0.8	5143	5	AF067418	AF067418 Xenopus l
841	18	0.8	1500	8	AF168909	Dimerocos	914	18	0.8	5714	2	AC015032	AC015032 Drosophil
842	18	0.8	1500	8	AF168930	Monococu	915	18	0.8	6608	1	AB070346	AB070346 Tetragono
843	18	0.8	1500	8	AF168949	Tapeinoc	916	18	0.8	6763	6	BD004101	BD004101 Apoptosis
844	18	0.8	1500	8	AF168950	Tradescan	917	18	0.8	6927	9	AB011103	AB011103 Mus muscu
845	18	0.8	1536	6	ARI38293	Sequence	918	18	0.8	7098	2	AC023900	AC023900 Homo sapi
846	18	0.8	1536	9	HSTIF	X79538 H. sapiens n	919	18	0.8	7292	8	CAERCTF	CAERCTF Mus muscu
847	18	0.8	1553	9	HUMFGFRB	L78720 Human fibro	920	18	0.8	7403	1	STU68493	STU68493 Candida alb
848	18	0.8	1581	6	AX489060	Sequence	921	18	0.8	7582	14	AF492770	AF492770 Sindbis v
849	18	0.8	1608	9	AF285158	AF285158 Homo sapi	922	18	0.8	8414	9	AY038064	AY038064 Homo sapi
850	18	0.8	1624	14	AB074067	AB074067 Homo sapi	923	18	0.8	8491	6	AX449276	AX449276 Sequence
851	18	0.8	1628	9	BC003662	BC003662 Homo sapi	924	18	0.8	8491	6	AX449276	AX449276 Sequence
852	18	0.8	1635	14	AB074069	AB074069 Homo sapi	925	18	0.8	8491	6	AX449278	AX449278 Sequence
853	18	0.8	1640	9	BC004386	BC004386 Homo sapi	926	18	0.8	8491	6	AX449278	AX449278 Sequence
854	18	0.8	1640	9	BC011151	BC011151 Homo sapi	927	18	0.8	8491	6	AX449278	AX449278 Sequence
855	18	0.8	1645	8	AF117298	AF117298 Homo sapi	928	18	0.8	8922	1	AB010235	AB010235 Pyrococcu
856	18	0.8	1648	14	AB074068	AB074068 Homo sapi	929	18	0.8	9099	8	AF315315	AF315315 Nectria h
857	18	0.8	1674	10	MMU46690	MMU46690 Human imm	930	18	0.8	9190	14	HIV0YI	HIV0YI Human immu
858	18	0.8	1682	6	ARI38294	Sequence	931	18	0.8	9771	1	U67492	U67492 Methanococc
859	18	0.8	1682	9	HUMORFJA	ARI38294 Sequence	932	18	0.8	10029	1	AE006797	AE006797 Sulfolobu
860	18	0.8	1748	3	DDHER2A	DDHER2A Human mRNA	933	18	0.8	10039	1	AE013108	AE013108 Thermomna
861	18	0.8	1779	5	GGG223618	AJ23618 D. discoiden	934	18	0.8	10082	4	AY071822	AY071822 Bos tauru
862	18	0.8	1780	10	MMU09138	MMU09138 Gallus ga	935	18	0.8	10137	1	AE013296	AE013296 Methanosa
863	18	0.8	1780	10	BC021798	BC021798 Mus muscu	936	18	0.8	10183	1	AE007790	AE007790 Clostridi
864	18	0.8	1800	10	U01664	U01664 Mus musculu	937	18	0.8	10549	1	U67511	U67511 Methanococc
865	18	0.8	1800	10	U01664	U01664 Mus musculu	938	18	0.8	10549	1	U67511	U67511 Methanococc
866	18	0.8	1876	6	ARI59782	ARI59782 Sequence	939	18	0.8	10591	1	AE013379	AE013379 Methanosa
867	18	0.8	1876	6	BD010462	BD010462 SMAD7 and	940	18	0.8	11014	1	AE010530	AE010530 Fusobacte
868	18	0.8	1914	6	AR208574	AR208574 Sequence	941	18	0.8	11026	8	AB020753	AB020753 Arabidops



QY 361 CCAACACGGAAACAGAGGGTTTCCAGAGAGCGGCTATCGAGATGGAATTAATTCAGAA 420  
 Db 374 CCAACACGGAAACAGAGGGTTTCCAGAGAGCGGCTATCGAGATGGAATTAATTCAGAA 433  
 QY 421 GCTTCAGAGGCATACAGAGAGGTGGAAGAGGTATGTTCCAGAGTTCCTGAGAGATTT 480  
 Db 434 GCTTCAGAGGCATACAGAGAGGTGGAAGAGGTATGTTCCAGAGTTCCTGAGAGATTT 493  
 QY 481 GGTCTAGAGAGTCCAAATTAATGACTTAGACCCAGACGAAATGATGACCGGACGTGGC 540  
 Db 494 GGTCTAGAGAGTCCAAATTAATGACTTAGACCCAGACGAAATGATGACCGGACGTGGC 553  
 QY 541 CTTTGGTTCTAGAGAGCAGATATTAAGTGGCAAGGTAAATGATATCTTCAAGC 600  
 Db 554 CTTTGGTTCTAGAGAGCAGATATTAAGTGGCAAGGTAAATGATATCTTCAAGC 613  
 QY 601 AGAAGTGCAGTGGAGTGAACGAGGTGTTACAAAGTTTAAATGAAGAAATTAACA 660  
 Db 614 AGAAGTGCAGTGGAGTGAACGAGGTGTTACAAAGTTTAAATGAAGAAATTAACA 673  
 QY 661 GGCTCTGGAAAGAAATTCCTGGAAGTCAGAGACAGAGAGGAGAAAGTATGATCTAA 720  
 Db 674 GGCTCTGGAAAGAAATTCCTGGAAGTCAGAGACAGAGAGGAGAAAGTATGATCTAA 733  
 QY 721 GGACCAAAAGTGAACCTATACCCCTCTCCACCTGAGATGAGACTCCATCTTTGCA 780  
 Db 734 GGACCAAAAGTGAACCTATACCCCTCTCCACCTGAGATGAGACTCCATCTTTGCA 793  
 QY 781 CATATCAGACAGCATTAACCTTGACAAATACGACACTATTTCTTGGAAGTGTGGA 840  
 Db 794 CATATCAGACAGCATTAACCTTGACAAATACGACACTATTTCTTGGAAGTGTGGA 853  
 QY 841 CATATCAGACAGCATTAACCTTGACAAATACGACACTATTTCTTGGAAGTGTGGA 900  
 Db 854 CATATCAGACAGCATTAACCTTGACAAATACGACACTATTTCTTGGAAGTGTGGA 913  
 QY 901 AACAACTTGTAAAGCTGTTATTAAGTCTACTCTGTCGCAAAAATACAGTATTCCT 960  
 Db 914 AACAACTTGTAAAGCTGTTATTAAGTCTACTCTGTCGCAAAAATACAGTATTCCT 973  
 QY 961 ATCACTACTGACAGACGAGATTTGATGCTGCTCAACAGGGTCTGGGAAGACTGGC 1020  
 Db 974 ATCACTACTGACAGACGAGATTTGATGCTGCTCAACAGGGTCTGGGAAGACTGGC 1033  
 QY 1021 GCTTTCTCCCTAACAATTTTGGCTCATATGATGATGATGATGATGATGATGATGAT 1080  
 Db 1034 GCTTTCTCCCTAACAATTTTGGCTCATATGATGATGATGATGATGATGATGATGAT 1093  
 QY 1081 AAAGAGTTGACAGAACGAGATGATTAATTTAGACCAACCTCGAGAAATGATCAACGAG 1140  
 Db 1094 AAAGAGTTGACAGAACGAGATGATTAATTTAGACCAACCTCGAGAAATGATCAACGAG 1153  
 QY 1141 ATTATATTGGAAGCCAGAAAAATTTCTTTGGGACTGTGTAGAGCTGTTGTTATATAT 1200  
 Db 1154 ATTATATTGGAAGCCAGAAAAATTTCTTTGGGACTGTGTAGAGCTGTTGTTATATAT 1213  
 QY 1201 GGGGGAACCCGCTGGGACATTCATTTGACAAATAGTACAAAGCTGTAAATATATATGT 1260  
 Db 1214 GGGGGAACCCGCTGGGACATTCATTTGACAAATAGTACAAAGCTGTAAATATATATGT 1273  
 QY 1261 GCTACTCCTGGAAGACTGATGATATCATAGCAAAAGAAAGATTTGCTCAACAGATC 1320  
 Db 1274 GCTACTCCTGGAAGACTGATGATATCATAGCAAAAGAAAGATTTGCTCAACAGATC 1333  
 QY 1321 AAATACTAGTTTGTGATGAAGCTGATGCAATGTTGATATGAGTTTGTCCAGAAATG 1380  
 Db 1334 AAATACTAGTTTGTGATGAAGCTGATGCAATGTTGATATGAGTTTGTCCAGAAATG 1393  
 QY 1381 AAGAAGTTAATTTCTTGCCAGAGATGCAATCAAGAAACAGCGCAAAACCTTATGTTTC 1440  
 Db 1394 AAGAAGTTAATTTCTTGCCAGAGATGCAATCAAGAAACAGCGCAAAACCTTATGTTTC 1453

QY 1441 AGTGCACCTTTCCAGAGAAATTCAAAGGTGGCTGCAGAGTTTAAAGTCAAATTAAT 1500  
 Db 1454 AGTGCACCTTTCCAGAGAAATTCAAAGGTGGCTGCAGAGTTTAAAGTCAAATTAAT 1513  
 QY 1501 CTGTTTGTGCTGTGGAACAAGTGGTGGAGCATGTAGAGATGTTCAAGCAGACCGTTCTC 1560  
 Db 1514 CTGTTTGTGCTGTGGAACAAGTGGTGGAGCATGTAGAGATGTTCAAGCAGACCGTTCTC 1573  
 QY 1561 CAAGTTGGCCAGTCTCAAAAAGAGAAAGCTCGTTGAAATTTGCGCAAAACATAGGGGAT 1620  
 Db 1574 CAAGTTGGCCAGTCTCAAAAAGAGAAAGCTCGTTGAAATTTGCGCAAAACATAGGGGAT 1633  
 QY 1621 GAAAGAACTATGATCTTTGTTGAAACTAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1680  
 Db 1634 GAAAGAACTATGATCTTTGTTGAAACTAGAAAAAAGCAGATTTTACTGCAACTTTTCTT 1693  
 QY 1681 TGTCAAGAAAAAATATCACTACAGATATCCATGTGTATCGGGAACAGAGAGCGGGAG 1740  
 Db 1694 TGTCAAGAAAAAATATCACTACAGATATCCATGTGTATCGGGAACAGAGAGCGGGAG 1753  
 QY 1741 CAAGCTCTGGAGATTTTGGCTTGGAAAGGCCAGTCTGTTGCTACTCTAGTACT 1800  
 Db 1754 CAAGCTCTGGAGATTTTGGCTTGGAAAGGCCAGTCTGTTGCTACTCTAGTACT 1813  
 QY 1801 GCCAGAGGCGTGAATATTGAAATGTGCAACATGTTATCAATTTTGAATCTTCTTACC 1860  
 Db 1814 GCCAGAGGCGTGAATATTGAAATGTGCAACATGTTATCAATTTTGAATCTTCTTACC 1873  
 QY 1861 ATTGATGAATATGTTCAATCGAATTTGGCGCTACTGCTGTGTGGGAATACTGCGAGACA 1920  
 Db 1874 ATTGATGAATATGTTCAATCGAATTTGGCGCTACTGCTGTGTGGGAATACTGCGAGACA 1933  
 QY 1921 ATTTCCTTTTGTGATCTTGAATGGATACCACTTTTACACAGCCTCTAGTAAAGTATG 1980  
 Db 1934 ATTTCCTTTTGTGATCTTGAATGGATACCACTTTTACACAGCCTCTAGTAAAGTATG 1993  
 QY 1981 ACAGATGCTCAACAGATGTTCTCGATGTTGGAAAGAAATGTCCTTAGATACATTAAT 2040  
 Db 1994 ACAGATGCTCAACAGATGTTCTCGATGTTGGAAAGAAATGTCCTTAGATACATTAAT 2053  
 QY 2041 CTGCGCTTCAGTGTAGTACAGAGAAAGCTGTTTCATCATGTTGATACCAAGAAAGGCG 2100  
 Db 2054 CTGCGCTTCAGTGTAGTACAGAGAAAGCTGTTTCATCATGTTGATACCAAGAAAGGCG 2113  
 QY 2101 AAGGACCTTGAACACAGCGGCTTTCTTCTTCAAGAGCTCCCAATCCAGTAGATGAT 2160  
 Db 2114 AAGGACCTTGAACACAGCGGCTTTCTTCTTCAAGAGCTCCCAATCCAGTAGATGAT 2173  
 QY 2161 GAGTCATGGAT 2172  
 Db 2174 GAGTCATGGAT 2185

RESULT 2  
 AF262962 2411 bp mRNA linear PRI 30-MAY-2000  
 LOCUS  
 DEFINITION Homo sapiens VASA protein mRNA, complete cds.  
 ACCESSION AF262962  
 VERSION AF262962.1 GI:8102020  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 2411)  
 Roeha, D. and Alfara, N.  
 Cloning and characterization of the human VASA gene  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 2411)  
 Roeha, D. and Alfara, N.  
 Direct Submission  
 Submitted (02-MAY-2000) Department of Pathology, University of  
 Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

FEATURES  
source

CDS

Location/Qualifiers  
1..2411  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
18..2192  
/codon\_start=1  
/product="VISA protein"  
/protein\_id="AAF72705.1"  
/db\_xref="GI:8102021"  
/translation="MGEDWEAEINPHMSVVPFEKDRISGENGDNPNRPASSSEM  
DDGSRDHFMSKSGFASGRNFGNRDAGCNDRNTSTVGGVKSKFNRFSRRE  
DDSSGFWRSSNDCEPDPTNRGSGKSDRDNNSGAPYRRGGYRGRGGF  
GLGSPNDLDDECMORTGLFGSRRPVLSTGNDPTSQSRSGSGSEGGYKLNBEV  
ITSGKNSKSEAEGBSSDPDQSPKVTYIPPEPDEDSIFAHYOTGINPDKXDTILY  
EVSGHDAPALITFEENALCOTLNNINKAKYTKLTPQKSTIPIIAGRPMACOT  
GSGKTAFLPLILAHMMHDGITASFKELOPECTIVAPTELVNQIYLEKRSFGT  
CVRVAVIYGQTQLSHSIRQIVQGNILCATGRLMDIIGKEKIGLKQIKYLVDEADR  
MLDMGPEMKKLISCPGMSRQROTILFSGTPEELIORSATPELISNYPFVAVGV  
GACRDVQOETVLOVGQPSKREKLEIILNIDERTMVEVTEKKADFLATPLCEKES  
TTSIHGDEOREREOAGDPRPKPCPVIVATSVARGLDIENVQINPDLPTIDY  
VHRIGTRRCENTRAISFPLBESDNHQAQVLVYLTDAODVPAMLEELAFSTYIEG  
FSGSTRGNVFASVDFRKGKSTLNTIRGESSQAENPVDESMD"

BASE COUNT 771 a 408 c 571 g 661 t  
ORIGIN

Query Match 91.1%; Score 1978; DB 9; Length 2411;  
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 3; Indels 0; Gaps 0;  
Matches 2128; Conservative 0;

7 GATGAAAGTTGGAGACAGAAATCAACCTCATATGCTCTCATGTTCCCATATTGAG 66  
24 GATGAAAGTTGGAGACAGAAATCAACCTCATATGCTCTCATGTTCCCATATTGAG 83  
67 AAGGATAGTATTTCTGAGAAAATGAGACAAATTTTAAACAGACTCCAGCTTCATATCA 126  
84 AAGGATAGTATTTCTGAGAAAATGAGACAAATTTTAAACAGACTCCAGCTTCATATCA 143  
127 GAAATGATGATGAGCTTTCTGAGAGATCATTTTATGAAAAGTGATTTGGCTCTGGG 186  
144 GAAATGATGATGAGCTTTCTGAGAGATCATTTTATGAAAAGTGATTTGGCTCTGGG 203  
187 CGGAATTTTGGAAACAGAGATGCTGTGATGTAATTAAGCGATATATATCCACAATG 246  
204 CGGAATTTTGGAAACAGAGATGCTGTGATGTAATTAAGCGATATATATCCACAATG 263  
247 GGTGTTTTGAGATTGAGAAAGATTTTGGAAAACAGAGTTTTTCAAACAGCAGTTTGA 306  
264 GGTGTTTTGAGATTGAGAAAGATTTTGGAAAACAGAGTTTTTCAAACAGCAGTTTGA 323  
307 GATGATGATGATGCTGTGTTCTGAGAGATGCTATGATTAATGCTGGAAGATATATCCACA 366  
324 GATGATGATGATGCTGTGTTCTGAGAGATGCTATGATTAATGCTGGAAGATATATCCACA 383  
367 CGGAACAGAGGTTTTTCCAAAGAGCGGCTATCGAGATGGAATTAATTTAGAAAGCTTCA 426  
384 CGGAACAGAGGTTTTTCCAAAGAGCGGCTATCGAGATGGAATTAATTTAGAAAGCTTCA 443  
427 GGGCCATACAGAAAGGTGAGAGAGTATTTCCAGAGTTGCGTGAGAGATTTGGTCTA 486  
444 GGGCCATACAGAAAGGTGAGAGAGTATTTCCAGAGTTGCGTGAGAGATTTGGTCTA 503  
487 GGAATTTCCAAATAATAGACTTGAAGCCAGAAATGATAGCAGCGGACGTGGGCTTTT 546  
504 GGAATTTCCAAATAATAGACTTGAAGCCAGAAATGATAGCAGCGGACGTGGGCTTTT 563  
547 GATTCTAGAGACAGATTTAATGAGCAGAGTATGATATCTTCAAAGCAGAGT 606  
564 GATTCTAGAGACAGATTTAATGAGCAGAGTATGATATCTTCAAAGCAGAGT 623  
607 GGCAGTGAAGTGAACGAGTGTGTTAAAGGTTTTAAATGAAGATATATACAGGCTCT 666  
624 GGCAGTGAAGTGAACGAGTGTGTTAAAGGTTTTAAATGAAGATATATACAGGCTCT 683

667 GGAAGAATTTCTGGAAGTCAGAGACAGAGAGGAAAGTAGTATCTCAAGACCA 726  
684 GGAAGAATTTCTGGAAGTCAGAGACAGAGAGGAAAGTAGTATCTCAAGACCA 743  
727 AAGTGACCTTACATACCCCTCTCTCCAGCTGAGATGAGACATCTTTGACATTA 786  
744 AAGTGACCTTACATACCCCTCTCTCCAGCTGAGATGAGACATCTTTGACATTA 803  
787 CAGACAGGCAATTAACCTTCCGAAATAGCACTATCTTGGAGGTCTGGACATGAT 846  
804 CAGACAGGCAATTAACCTTCCGAAATAGCACTATCTTGGAGGTCTGGACATGAT 863  
847 GCACCAACCAATTTCTGACTTTTGAAGAAGTAACTCTGTGACACACTGAATTAAC 906  
864 GCACCAACCAATTTCTGACTTTTGAAGAAGTAACTCTGTGACACACTGAATTAAC 923  
907 ATTGCTAAAGCTGTTATTAAGCTTACTCTGTGCAAAATACAGTATTTCTATCAT 966  
924 ATTGCTAAAGCTGTTATTAAGCTTACTCTGTGCAAAATACAGTATTTCTATCAT 983  
967 CTTCGAGAGAGATTTGATGCGCTTGCTCAACAGGCTCTGGAAAGACTGGGGCTTT 1026  
984 CTTCGAGAGAGATTTGATGCGCTTGCTCAACAGGCTCTGGAAAGACTGGGGCTTT 1043  
1027 CTCTCAACCAATTTTGGCTCATATGATGATGATGATTAACCTGCACTGTTTTAAAG 1086  
1044 CTCTCAACCAATTTTGGCTCATATGATGATGATGATTAACCTGCACTGTTTTAAAG 1103  
1087 TTGAGAGAACCAAGATGATTAATTTGTCACCAACTGAGAAATTTGTCACCAAGT 1146  
1104 TTGAGAGAACCAAGATGATTAATTTGTCACCAACTGAGAAATTTGTCACCAAGT 1163  
1147 TTGAGAACCAAGAAATTTTCTTTTGGGACTTGCTGTAAGCTGTTGTAATATGAGG 1206  
1164 TTGAGAACCAAGAAATTTTCTTTTGGGACTTGCTGTAAGCTGTTGTAATATGAGG 1223  
1207 ACCAGCTGGGACATTTCAATTTCCGCAATATGATCAAGCTGTAATATATGCTACT 1266  
1224 ACCAGCTGGGACATTTCAATTTCCGCAATATGATCAAGCTGTAATATATGCTACT 1283  
1267 CTTGGAAGACTGATGATATCATAGGCAAAAGAAAGTTGCTCTCAACAGATCAATAT 1326  
1284 CTTGGAAGACTGATGATATCATAGGCAAAAGAAAGTTGCTCTCAACAGATCAATAT 1343  
1327 TTATGTTTTGATGAGAGCTGATCGCATTTGGATATGAGTTTGGTCCAGAAATGAAG 1386  
1344 TTATGTTTTGATGAGAGCTGATCGCATTTGGATATGAGTTTGGTCCAGAAATGAAG 1403  
1387 TTATGTTTTGATGAGAGCTGATCGCATTTGGATATGAGTTTGGTCCAGAAATGAAG 1446  
1404 TTATGTTTTGATGAGAGCTGATCGCATTTGGATATGAGTTTGGTCCAGAAATGAAG 1463  
1447 ACTTTTCCAGAGAAATTAAGAGTTGGCTGACAGATTTTAAAGTCAATATATCTGTT 1506  
1464 ACTTTTCCAGAGAAATTAAGAGTTGGCTGACAGATTTTAAAGTCAATATATCTGTT 1523  
1507 GTTCTGTTGACAAAGTGGGTGAGCATGTAGAGATTTTCAAGCAAGCCGTTCTCAAGT 1566  
1524 GTTCTGTTGACAAAGTGGGTGAGCATGTAGAGATTTTCAAGCAAGCCGTTCTCAAGT 1583  
1567 GGCAGTTCTCAAAAAGAGAAAGCTGTTGAAATTTCTGCAAACTATGAGGATGAAGA 1626  
1584 GGCAGTTCTCAAAAAGAGAAAGCTGTTGAAATTTCTGCAAACTATGAGGATGAAGA 1643  
1627 ACTATGCTCTTGTGTAAGCACTTAAGAAAAAGACAGATTTTACTGCACTTTTGTCAA 1686  
1644 ACTATGCTCTTGTGTAAGCACTTAAGAAAAAGACAGATTTTACTGCACTTTTGTCAA 1703  
1687 GAAAAAATATCACTTAAGATTCATGATGATGAGGAAACAGAGAGCGGAGCAAGCT 1746  
1704 GAAAAAATATCACTTAAGATTCATGATGATGAGGAAACAGAGAGCGGAGCAAGCT 1763

QY 1747 CTGGAGATTTTCGCTTGGAAAGTCCAGTCTTGTGCTACTTCACTGACTGACCA 1806  
DB 1764 CTGGAGATTTTCGCTTGGAAAGTCCAGTCTTGTGCTACTTCACTGACTGACCA 1823  
QY 1807 GGGCTGATTTGAATAATGTGCAACATGTTATCAATTTGATCTTCTCTCACTGAT 1866  
DB 1824 GGGCTGATTTGAATAATGTGCAACATGTTATCAATTTGATCTTCTCTCACTGAT 1883  
QY 1867 GAATATGTTCAATCGAATTTGGGCGTACTGCTGCTTGGGAATACTGGAGCAATTTCC 1926  
DB 1884 GAATATGTTCAATCGAATTTGGGCGTACTGCTGCTTGGGAATACTGGAGCAATTTCC 1943  
QY 1927 TTTTGTATCTTGAATTCGATATACATTTAGACACGCTCTAGTAATAAGTATGACAT 1986  
DB 1944 TTTTGTATCTTGAATTCGATATACATTTAGACACGCTCTAGTAATAAGTATGACAT 2003  
QY 1987 GCTCAACAGATTTGCTTCTGCAATGTTGGAAGAAATGCTTATGACATCTTCTGAC 2046  
DB 2004 GCTCAACAGATTTGCTTCTGCAATGTTGGAAGAAATGCTTATGACATCTTCTGAC 2063  
QY 2047 TTCAGTGTAGTACAAAGAGAAACGTGTTGCTATGATGATACAGAAAGGCAAGAGC 2106  
DB 2064 TTCAGTGTAGTACAAAGAGAAACGTGTTGCTATGATGATGATACAGAAAGGCAAGAGC 2123  
QY 2107 ACTTGAACACAGCTGGGTTTCTTCTTCTAC 2137  
DB 2124 ACTTGAACACAGCTGGGTTTCTTCTTCTAC 2154

RESULT 3  
HSM802178 2189 bp mRNA linear PRI 18-FEB-2000  
LOCUS Homo sapiens mRNA; cDNA DKFZp434B1122 (from clone DKFZp434B1122);  
DEFINITION partial cds.  
ACCESSION AL137462  
VERSION AL137462.1 GI:6808043  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J., and Wiemann, S.  
TITLE Direct Submission  
JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Sequenced by LMU (Ludwig Maximilians University,  
Munich/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
This clone (DKFZp434B1122) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
information about the clone and the sequencing project is available  
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.  
FEATURES  
source  
1..2189  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZp434B1122"  
/tissue\_type="testis"  
/clone\_lib="434 (synonym: htes3). Vector pSport1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
1..1910  
/gene="DKFZp434B1122"  
/gene="DKFZp434B1122"  
/note="strong similarity to testis specific Helicase (VASA  
homolog)"  
/codon\_start=3  
/product="hypothetical protein"

/protein\_id="CAB70750.1"  
/db\_xref="gi:6808044"  
/translation="SFGNRGFSNRFEDGSSGCFWRBSSNDCEPNPTNRGFSKRGY  
RDGNNSEASGPRRGSGSPRCGRGGGGLGSPNNLDLDECMORGTGLFGSRRLVPLSP  
TGNDTDSRGSGSSEBGGYKGLNEUYTSGSKNSKMSKSEAGGSSSDTQGRVYIIP  
PPDEDISIFAHYOTGTFNFDKDTTLIVESGSDNPALLITFEANLCTLNNLAKAG  
YTKLTPVQKRSIPILLAGRDLMAAGQSGSTAAFLPLILAHMHMDITJAFRFELOE  
PECTIAPTRRELNVQIYLEAKFSFGCVRAVVIYGGTOLGHSIROIVGCNLTATP  
GRLMIDIGKEKIGLKEIYGLADPADMLDMGFPPEMKLISCGMPSKEORLTMS  
ATPEEIORLAELFKSNLYLFAVAVGVGAARDVQOYLVQVGSFKPKVLEIRNTG  
DERIMVETKKAADFIATFLFCQKISTTISHGRBQREORALGDFRFGCPVLVAT  
SVARGLDIEVQHVINFDPSTLIDEVVHRIGRGRCGNTGRLAISFDLESNDHLAOP  
LVKVLTAQODVPVPMLEIAFSTYIPGSGSTRGNVPAVDTRKSKTLNTAGSSSQ  
APNVDDESMD"  
polyA\_signal  
polyA\_site  
BASE COUNT 2128. 2133  
polyA 706 a 373 c 505 g 605 t  
ORIGIN

Query Match 79.1%; Score 1719; DB 9; Length 2189;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1869; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 266 AGAGTTTGGAAACAGAGCTTTTCAACACAGAGTTGAAGATGTGATAGCTTGCTT 325  
DB 1 AGAGTTTGGAAACAGAGCTTTTCAACACAGAGTTGAAGATGTGATAGCTTGCTT 60  
QY 326 TCTGAGAGAGTCTAGTAATGACATCGCAAGATATCCAAACGGAACAGAGGTTTCCA 385  
DB 61 TCTGAGAGAGTCTAGTAATGACATCGCAAGATATCCAAACGGAACAGAGGTTTCCA 120  
QY 386 AGAGAGGCGCTATCGAGATGGAATAATTCAGAGCTTCAGAGCCATACAGAGAGGTG 445  
DB 121 AGAGAGGCGCTATCGAGATGGAATAATTCAGAGCTTCAGAGCCATACAGAGAGGTG 180  
QY 446 GAAAGGTATGTTCCGAGGTGCCGTGAGAGATTTGCTAGAGATCCAAATAATGACT 505  
DB 181 GAAAGGTATGTTCCGAGGTGCCGTGAGAGATTTGCTAGAGATCCAAATAATGACT 240  
QY 506 TAGACCCAGAGATGTAATGAGAGGCACTGGTGGCTTTTGTGTTCTAGAGACCAAT 565  
DB 241 TAGACCCAGAGATGTAATGAGAGGCACTGGTGGCTTTTGTGTTCTAGAGACCAAT 300  
QY 566 TAGGTGCAAGATGTAATGAGATGTAATGAGATGTAATGAGATGTAATGAGATGTA 625  
DB 301 TAGGTGCAAGATGTAATGAGATGTAATGAGATGTAATGAGATGTAATGAGATGTA 360  
QY 626 GTGTTTCAAAAGGTTTAAATGAAAGTAAATGAAAGTAAATGAAAGTAAATGAAAGT 685  
DB 361 GTGTTTCAAAAGGTTTAAATGAAAGTAAATGAAAGTAAATGAAAGTAAATGAAAGT 420  
QY 686 CAGAGCAG 745  
DB 421 CAGAGCAG 480  
QY 746 CTCTCCACCTGAGAGATGAGAGCTCCATCTTTGACATTTATCAACAGGCTAAACTTCG 805  
DB 481 CTCTCCACCTGAGAGATGAGAGCTCCATCTTTGACATTTATCAACAGGCTAAACTTCG 540  
QY 806 ACAATAACGACATATTTCTGTGAAAGTGTCTGACATGATGACCAACAGCAATTCG 865  
DB 541 ACAATAACGACATATTTCTGTGAAAGTGTCTGACATGATGACCAACAGCAATTCG 600  
QY 866 CTTTGAAGAGATTAATCTCTGTAGACATGTAATTAACATTTGCTTAAGCTGTTATA 925  
DB 601 CTTTGAAGAGATTAATCTCTGTAGACATGTAATTAACATTTGCTTAAGCTGTTATA 660  
QY 926 CTAAGCTTACCTCTGTGCAAAATATAGATATCTTATCATATCTGACGAGAGATTTGA 985  
DB 661 CTAAGCTTACCTCTGTGCAAAATATAGATATCTTATCATATCTGACGAGAGATTTGA 720  
QY 986 TGCGTTGTCTCAAAAGGCTTGGAGAACTGGGCTTTTCTCTCAATTTTGCTC 1045



Db	721	IGGCTTGGCCCTCAAAACAGGGCTCGGGAAGAATCGGGCTTTCTCTCAACCAATTTTGGCTC	780
Qy	1046	ATATGATGCATGATGGAATTAATCCCACTCGTTTAAAGAGTGCAGAAACAGAGTGA	1105
Db	781	ATATGATGCATGATGGAATTAATCCCACTCGTTTAAAGAGTGCAGAAACAGAGTGA	840
Qy	1106	TTATTTGAGCAACCACTGAGAAATTTGGTCAACCAATTTATTGGAAGCCAGAAATTTT	1165
Db	841	TTATTTGAGCAACCACTGAGAAATTTGGTCAACCAATTTATTGGAAGCCAGAAATTTT	900
Qy	1166	CTTTTGGGACCTTGCTAAGAGCTGTTGTTATATATGGGGGAACCCACATGGGACATTCAA	1225
Db	901	CTTTTGGGACCTTGCTAAGAGCTGTTGTTATATATGGGGGAACCCACATGGGACATTCAA	960
Qy	1226	TTGCAAAATAGTACAAAGCTGTATATATATATATGTTACTCTGTGAAGACTGATGA	1285
Db	961	TTGCAAAATAGTACAAAGCTGTATATATATATATGTTACTCTGTGAAGACTGATGA	1020
Qy	1286	TCATTTGGCAAAAGAAAGATTTGGTCTCAACAGATCAAAATCTTAGTTTGGATGAGCTG	1345
Db	1021	TCATTTGGCAAAAGAAAGATTTGGTCTCAACAGATCAAAATCTTAGTTTGGATGAGCTG	1080
Qy	1346	ATCGCATGTGGATATGAGTTTGGTTCAGAAATGAAGAAAGTATATTTCTTGCCAGAA	1405
Db	1081	ATCGCATGTGGATATGAGTTTGGTTCAGAAATGAAGAAAGTATATTTCTTGCCAGAA	1140
Qy	1406	TGCCATCAAAAGGAACAGCGCCAAACCCCTATGTTCAGTCAAACTTTTCAAGAGAAATTC	1465
Db	1141	TGCCATCAAAAGGAACAGCGCCAAACCCCTATGTTCAGTCAAACTTTTCAAGAGAAATTC	1200
Qy	1466	AAAGGTGGCTGCAGAGTTTAAAGTCAAAATATATCTGTTGTTGCTGTTGACAAATGG	1525
Db	1201	AAAGGTGGCTGCAGAGTTTAAAGTCAAAATATATCTGTTGTTGCTGTTGACAAATGG	1260
Qy	1526	GTGGAGCACTGATGAGATGTCACGACGATTCGCAAGTTGGGACAGTTCAGAAAAG	1585
Db	1261	GTGGAGCACTGATGAGATGTCACGACGATTCGCAAGTTGGGACAGTTCAGAAAAG	1320
Qy	1586	AAAAGCTGTGAAATTTCTGCGAAACATAGGGAGTGAAGAAGTATGTTCTTTGTGAA	1645
Db	1321	AAAAGCTGTGAAATTTCTGCGAAACATAGGGAGTGAAGAAGTATGTTCTTTGTGAA	1380
Qy	1646	CTAAGAAAAAAGCAGATTTTACTGCACTTTTGTGCAGAAAAAATATCAACTCA	1705
Db	1381	CTAAGAAAAAAGCAGATTTTACTGCACTTTTGTGCAGAAAAAATATCAACTCA	1440
Qy	1706	GTATCCATGATGATGCGGGAACAAGAGACGGGAGCAAGCTCTTGAGATTTTGCTTG	1765
Db	1441	GTATCCATGATGATGCGGGAACAAGAGACGGGAGCAAGCTCTTGAGATTTTGCTTG	1500
Qy	1766	GAAATGCCCCAGTTCTTGTTGCTACTTACAGTAGCTGCCAGAGGGCTGATATGAAATG	1825
Db	1501	GAAATGCCCCAGTTCTTGTTGCTACTTACAGTAGCTGCCAGAGGGCTGATATGAAATG	1560
Qy	1826	TGCAACATGTTATCAATTTTGAATCTTCTTCAACATGATGAATGTTCAATGAAATG	1885
Db	1561	TGCAACATGTTATCAATTTTGAATCTTCTTCAACATGATGAATGTTCAATGAAATG	1620
Qy	1886	GGCGTATCTGCTGTTGGGAAATACTGCGACAGCAATTTCTCTTTTGTGATCTTGAATCGG	1945
Db	1621	GGCGTATCTGCTGTTGGGAAATACTGCGACAGCAATTTCTCTTTTGTGATCTTGAATCGG	1680
Qy	1946	ATTAACATTTAGCACAGCTCTAGTATAAAGTATATACAGATGCTCAACAGGATGTTCTG	2005
Db	1681	ATTAACATTTAGCACAGCTCTAGTATAAAGTATATACAGATGCTCAACAGGATGTTCTG	1740
Qy	2006	CATGTTGGAGAAATTTGCTTTAGTACATATCTTGCTTCAGTGTGATCAAGAG	2065
Db	1741	CATGTTGGAGAAATTTGCTTTAGTACATATCTTGCTTCAGTGTGATCAAGAG	1800
Qy	2066	GAAAGCTGTTTGATCATGTTGATATACCAAGAAAGGCAAGACATTTGAACACAGTGGT	2125
Db	1801	GAAAGCTGTTTGATCATGTTGATATACCAAGAAAGGCAAGACATTTGAACACAGTGGT	1860

QY	2126	TTTCTCTTCAC	2137																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
----	------	-------------	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



OY		2040	TCCGGGCTTCAAGTGATGTACAAAGAAAAGCGTTTGACATCAGTTATTCAGAAAGGG	2099
Db		1612	TCTGGCTTCAGTGTAAGTAGACAGAAGAACGTTTTGCATCATCTGATACCAAGAGGG	1671
OY		2100	CAGAGCACCTTTGAACAAGCTGGGTTTTCTTCTTCAC	2137
Db		1672	CAGAGCACCTTTGAACAAGCTGGGTTTTCTTCTTCAC	1709
RESULT 5				
LOCUS	AC008914	139677 bp	DNA	linear PRI 19-APR-2002
DEFINITION	Homo sapiens chromosome 5 clone CTD-2270116, complete sequence.			
ACCESSION	AC008914			
VERSION	AC008914.7			
KEYWORDS	HTG.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
JOURNAL	DOE Joint Genome Institute and Stanford Human Genome Center.			
REFERENCE	1 (bases 1 to 139677)			
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.			
JOURNAL	Direct Submission			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 139677)			
JOURNAL	DOE Joint Genome Institute.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint			
JOURNAL	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
REFERENCE	3 (bases 1 to 139677)			
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (27-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell			
AUTHORS	Drive, Walnut Creek, CA 94598, USA			
JOURNAL	4 (bases 1 to 139677)			
REFERENCE	DOE Joint Genome Institute and Stanford Human Genome Center.			
AUTHORS	Submitted (19-APR-2002) DOE Joint Genome Institute, 2800 Mitchell			
JOURNAL	Drive, Walnut Creek, CA 94598, USA			
COMMENT	On Apr 19, 2002 this sequence version replaced gi:19747123.			
COMMENT	Draft Sequence Produced by DOE Joint Genome Institute			
COMMENT	www.jgi.doe.gov			
COMMENT	Finishing Completed at Stanford Human Genome Center			
COMMENT	Quality: Phrap Quality >=40 99.8% of Sequence;			
COMMENT	Estimated Total Number of Errors is 0.2.			
FEATURES	Location/Qualifiers			
source	1..139677			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="5"			
	/clone="CTD-2270116"			
BASE COUNT	37913 a 28849 c 28939 g 43976 t			
ORIGIN				
Query Match	12.6%; Score 273; DB 9; Length 139677;			
Best Local Similarity	100.0%; Pident.No. 5.8e+40; Indels 0; Gaps 0;			
Matches 273; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
OY		1715	GTCATCGGAAACAGAGAGAGCGGAGCAGACTCTTGAGATTTCGCTTGGAAAGGCC	1774
Db		13765	GTCATCGGAAACAGAGAGAGCGGAGCAGACTCTTGAGATTTCGCTTGGAAAGGCC	1382
OY		1775	CAGTCTCTTGCTACTTCAGTAGCTCCACAGAGGCTGATATTGAAATGTGCACATG	1834
Db		13825	CAGTCTCTTGCTACTTCAGTAGCTCCACAGAGGCTGATATTGAAATGTGCACATG	1388
OY		1835	TTATCAATTTGATCTTCCTTCAACATGATGAATAATGTTCAATCGAATTGGGGCTACTG	1894
Db		13885	TTATCAATTTGATCTTCCTTCAACATGATGAATAATGTTCAATCGAATTGGGGCTACTG	1394
OY		1895	GTCCTTGCGGAATCTGCGAGACAACTTCCTTTTGTAGCTTGAATCGATACCATT	1954

Db 13945 GTCCGTTGGGAATACGTGCAGAGCAATTCCTTTTGTGATCTTGATCGATAACCACTT 14004  
 QY 1955 TAGCAGACCTCTAGTAAAGATGACAGATG 1987  
 Db 14005 TAGCAGACCTCTAGTAAAGATGACAGATG 14037

RESULT 6  
 AC016632/c  
 LOCUS Homo sapiens chromosome 5, clone RP11-175M2, WORKING DRAFT SEQUENCE,  
 DEFINITION 1 unordered piece.  
 AC016632  
 AC016632.6 GI:15290348  
 VERSION HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS Homo sapiens.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 176784)  
 TITLE DOE Joint Genome Institute.  
 AUTHORS Sequencing of Human Chromosome 5  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 176784)  
 TITLE DOE Joint Genome Institute.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Aug 25, 2001 this sequence version replaced gi:13699595.  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: http://www.jgi.doe.gov  
 -----  
 Project Information  
 Center Project Name: 473855  
 Center clone name: RPCI-11\_175M2  
 -----  
 Summary Statistics  
 Consensus quality: 175566 bases at least Q40  
 Consensus quality: 176782 bases at least Q30  
 Consensus quality: 176784 bases at least Q20  
 Estimated insert size: 182000; pulse field gel estimation  
 Estimated insert size: 176784; sum-of-contigs estimation  
 Quality coverage: 7.55 in Q20 bases; pulse field gel estimation  
 Quality coverage: 7.78 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 176784: contig of 176784 bp in length.  
 Location/Qualifiers  
 1. 176784  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="RP11-175M2"  
 /clone\_1fb="RPCI human BAC library 11"  
 BASE COUNT 57392 a 35064 c 33684 g 50644 t  
 ORIGIN

Query Match 12.6%; Score 273; DB 2; Length 176784;  
 Best Local Similarity 100.0%; Pred.No. 5.7e-140;  
 Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1715 GTGATCGGGAACAGAGAGCGGAGCAAGCTCTTGAGATTTTGGCTTTGGAAAGTCC 1774  
 Db 34889 GTGATCGGGAACAGAGAGCGGAGCAAGCTCTTGAGATTTTGGAAAGTCC 34830

QY 1775 CAGTTCTTGTGCTACTTCACTAGTCCGACAGGGCTGATATTGAATGTGCACATG 1834  
 Db 34829 CAGTTCTTGTGCTACTTCACTAGTCCGACAGGGCTGATATTGAATGTGCACATG 34770  
 QY 1835 TTATCAATTTTGATCTTCTCTTCAACATTGATGATTTGTCATCGAATGGGCTACTG 1894  
 Db 34769 TTATCAATTTTGATCTTCTCTTCAACATTGATGATTTGTCATCGAATGGGCTACTG 34710

QY 1895 GTGCTGTGGGAATACGTGCAGAGCAATTCCTTTTGTGATCTTGATCGGATTAACCAT 1954  
 Db 34709 GTGCTGTGGGAATACGTGCAGAGCAATTCCTTTTGTGATCTTGATCGGATTAACCAT 34650

QY 1955 TAGCAGACCTCTAGTAAAGATTTGACAGATG 1987  
 Db 34649 TAGCAGACCTCTAGTAAAGATTTGACAGATG 34617

RESULT 7  
 AC016639/c  
 LOCUS Homo sapiens chromosome 5 clone RP11-412L4, WORKING DRAFT SEQUENCE,  
 DEFINITION 7 ordered pieces.  
 AC016639  
 AC016639.5 GI:7711585  
 VERSION HTG; HTGS PHASE2; HTGS\_DRAFT.  
 KEYWORDS Homo sapiens.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 182126)  
 TITLE DOE Joint Genome Institute.  
 AUTHORS Sequencing of Human Chromosome 5  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 182126)  
 TITLE DOE Joint Genome Institute.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On May 6, 2000 this sequence version replaced gi:7710162.  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: http://www.jgi.doe.gov  
 -----  
 Project Information  
 Center Project Name: 564841  
 Center clone name: RPCI-11\_412L4  
 -----  
 Summary Statistics  
 Consensus quality: 175214 bases at least Q40  
 Consensus quality: 180404 bases at least Q30  
 Consensus quality: 181112 bases at least Q20  
 Estimated insert size: 178000; pulse field gel estimation  
 Estimated insert size: 181876; sum-of-contigs estimation  
 Quality coverage: 6.41 in Q20 bases; pulse field gel estimation  
 Quality coverage: 6.28 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 48575: contig of 48575 bp in length  
 48576 48675: gap of unknown length  
 48676 62810: contig of 14135 bp in length  
 62811 62910: gap of unknown length  
 62911 81913: contig of 19003 bp in length  
 81914 82013: gap of unknown length  
 82014 92166: contig of 10153 bp in length  
 92167 92266: gap of unknown length



```

----- Project Information -----
Center project name: H.NH0332C17
----- Summary Statistics -----
Sequencing vector: M13; 87%
Sequencing vector: plasmid; 13%
Chemistry: Dye-Primer ET; 85% of reads
Chemistry: Dye-Terminator Big Dye; 15% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196499 bases at least Q40
Consensus quality: 199226 bases at least Q40
Consensus quality: 200804 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 202150; sum-of-contigs
Quality coverage: 4.50 in Q20 bases; agarose-fp
Quality coverage: 4.65 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
-----
1      1062: contig of 1062 bp in length
*      1063      1162: gap of unknown length
*      1163      2536: contig of 1374 bp in length
*      2537      2636: gap of unknown length
*      2637      5791: contig of 3155 bp in length
*      5792      5891: gap of unknown length
*      5892      10946: contig of 5055 bp in length
*      10947      11046: gap of unknown length
*      11047      15725: contig of 4679 bp in length
*      15726      15825: gap of unknown length
*      15826      21169: contig of 5344 bp in length
*      21170      21693: gap of unknown length
*      21270      27800: contig of 6531 bp in length
*      27801      27800: gap of unknown length
*      27901      33276: contig of 5376 bp in length
*      33277      33576: gap of unknown length
*      33577      42711: contig of 9335 bp in length
*      42712      42811: gap of unknown length
*      42812      56046: contig of 13235 bp in length
*      56047      56146: gap of unknown length
*      56147      77263: contig of 21117 bp in length
*      77264      77363: gap of unknown length
*      77364      97974: contig of 20611 bp in length
*      97975      98074: gap of unknown length
*      98075      122372: contig of 24198 bp in length
*      122373      122372: gap of unknown length
*      122373      148199: contig of 25827 bp in length
*      148200      148299: gap of unknown length
*      148300      175035: contig of 26736 bp in length
*      175036      175135: gap of unknown length
*      175136      204250: contig of 29115 bp in length.
-----
FEATURES
Source
1..204250
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-332C17"
misc_feature
1..1062
/note="assembly_name:Contig12"
misc_feature
1163..2536
/note="assembly_name:Contig13"
misc_feature
2637..5791
/note="assembly_name:Contig14"
misc_feature
5892..10946
/note="assembly_name:Contig15"
misc_feature
11047..15725
/note="assembly_name:Contig16"
misc_feature
15826..21169
/note="assembly_name:Contig17"

```

```
misc_feature 21270..27800
/note="assembly_name:Contig18"
misc_feature 27901..33376
/note="assembly_name:Contig19"
misc_feature 33377..42711
/note="assembly_name:Contig20"
clone_end:17
vector_side:right"
misc_feature 42812..56046
/note="assembly_name:Contig21"
misc_feature 56147..77263
/note="assembly_name:Contig22"
misc_feature 77364..97974
/note="assembly_name:Contig23"
misc_feature 98075..122272
/note="assembly_name:Contig24"
misc_feature 122373..148199
/note="assembly_name:Contig25"
misc_feature 148300..175035
/note="assembly_name:Contig26"
misc_feature 175136..204250
/note="assembly_name:Contig27"
BASE COUNT 62979 a 41247 c 40075 g 58442 t 1503 others
ORIGIN
Query Match 8.0%; Score 174; DB 2; Length 204250;
Best Local Similarity 100.0%; Pred. No. 8.1e-85;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1299 AAAATTGGTCTCAACAGATCAAAATCTTGTGGATGAGCTGATGCCATGTTGCA 1358
Db 170192 AAAATTGGTCTCAACAGATCAAAATCTTGTGGATGAGCTGATGCCATGTTGCA 170133
QY 1359 TATGGGTTTGGTCAGAAATGAGAAGTATATTTCTGGCCAGGAATGCCATCAAGGA 1418
Db 170132 TATGGGTTTGGTCAGAAATGAGAAGTATATTTCTGGCCAGGAATGCCATCAAGGA 170073
QY 1419 ACAGCGCCAAACCCCTTATGTTCACTGCACTTTTCCAGAGAAATTCAAAGTT 1472
Db 170072 ACAGCGCCAAACCCCTTATGTTCACTGCACTTTTCCAGAGAAATTCAAAGTT 170019
RESULT 10
LOCUS AY100475 524 bp DNA linear MAM 15-JUN-2002
DEFINITION Equus caballus VASA-like protein gene, partial sequence.
ACCESSION AY100475
VERSION AY100475.1 GI:21429223
KEYWORDS
SOURCE horse.
ORGANISM Equus caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 524)
WOODS,B.G., Ginther,O.J., Wentworth,A., Wentworth,B. and
Wiltbank,M.
TITLE Equine VASA Homolog
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 524)
WOODS,B.G., Ginther,O.J., Wentworth,A., Wentworth,B. and
Wiltbank,M.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2002) Animal Health and Biomedical Sciences,
AUTHORS University of Wisconsin-Madison, 1656 Linden Drive, Madison, WI
53706, USA
FEATURES
source
1..524
Location/Qualifiers
misc_feature <1..524
/db_xref="taxon:9796"
BASE COUNT 153 a 94 c 115 g 151 t 11 others
ORIGIN
```

```
Query Match 2.7%; Score 58; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 4.9e-20;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1984 GATGCTCAACAGATGTTCTTCATGCTGTTGGAGAATAATGCTTTAGTACATACATTC 2041
Db 455 GATGCTCAACAGATGTTCTTCATGCTGTTGGAGAATAATGCTTTAGTACATACATTC 512
RESULT 11
LOCUS S75275 3030 bp mRNA linear ROD 15-MAY-1995
DEFINITION RVLG=vasa-like gene protein [rate, wistar-Imanishi, testis, mRNA,
3030 nt].
ACCESSION S75275
VERSION S75275.1 GI:806463
KEYWORDS
SOURCE Rattus sp. testis wistar-Imanishi.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3030)
AUTHORS Komiya,T. and Tanigawa,Y.
TITLE Cloning of a gene of the DEAD box protein family which is
specifically expressed in germ cells in rats
JOURNAL Biochem. Biophys. Res. Commun. 207 (1), 405-410 (1995)
MEDLINE 95160706
PUBMED 7857296
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gisbseq 159937] from the original journal article.
This sequence comes from Fig. 1.
FEATURES
source
1..3030
Location/Qualifiers
/organism="Rattus sp."
/db_xref="taxon:10118"
1..3030
/gene="RVLG"
/note="vasa-like gene protein, RVLG protein"
266..2407
/gene="RVLG"
/note="putative DEAD box protein; This sequence comes from
Fig. 1; RVLG protein"
/codon_start=1
/product="vasa-like gene protein"
/protein_id="AA83364.1"
/db_xref="GI:806464"
/translation="MGDEWEAEILKPHVSVYVPEKDKYSGANGDTFNRTSASS
EMEDPGSRDHPMRSGFSGRNIGNRDIGSSKRETTSTGGFGRGFGNRPDLNKK
FEQDSSGFPWKESTNDCEDPTOTRSRGSKGYPDGNDSSASGFRGCGDSVPDQD
GSQRGGLFSSRRKPAHSDSGSDTFQSSRNAGAKYGLNBEVVTGSGKMSKEABG
GESSDIQGPVTVYIPPPPPDEDSIFAHYGTGINFDYDTLIVESGDADAPALITTE
EANLCQTNNNNIAKAGTKLTPVKYSIPVLAGRDIMACQTSQTAFLPILAH
MMEDQIVASRFEKLOEPECIIVATRELINQIYLEAKFSFGTAVAVIYGTQFGH
SIRQIVOGCNIILCATPGRIMDIIGKEKIGIKOYVYLDEADRLMDMGFGEMKLI
CPGMSKEOROTLFSATPPEIOTRAGLEKATLGVAVAGVCGACRDNOOSTIYNG
PVKKRKVLIELNIGDERPMVETVEYTKKADFLATFLCQEKISTTSIHGBQREBO
ALDGFRCGCPVLAVTSVAARGLDIENVQVINENLSTIDEVYHRIGRGRGCGNTR
AISFPDTSNHNLAQPLKVLSDAQDVPAMLEBIARSSVAPFSFNSGTGAVAFSPD
TRKNFGKNTLNTAGISSAQAHPNVDDESD"
BASE COUNT 900 a 551 c 723 g 856 t
ORIGIN
Query Match 2.4%; Score 53; DB 10; Length 3030;
Best Local Similarity 100.0%; Pred. No. 2.7e-17;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1141 ATTATTGGAAGCCAGAAATTTCTTTGGGAGCTTGTAAGAGCTGTGT 1193
Db 1361 ATTATTGGAAGCCAGAAATTTCTTTGGGAGCTTGTAAGAGCTGTGT 1413
```

RESULT 12  
LOCUS MUSDVH 1930 bp mRNA linear ROD 04-FEB-1999  
DEFINITION Mouse mRNA for drosophila vasa homologue, partial cds.  
ACCESSION D14859.1 GI:286074  
VERSION D14859.1  
KEYWORDS RNA helicase; drosophila vasa homologue.  
SOURCE Mus musculus (strain BALB/c) (library: lambda gt10) adult and embryo gonad and testis primordial germ cell, spermatogonium and spermatocyte cDNA to mRNA, clone mvr.  
ORGANISM Mus musculus  
REFERENCE 1 (sites)  
AUTHORS Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M. and Noce, T.  
TITLE Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosophila vasa and its specific expression in germ cell lineage  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (25), 12258-12262 (1994)  
MEDLINE 95083681  
REFERENCE 2 (bases 1 to 1930)  
AUTHORS Noce, T.  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 1930)  
AUTHORS Noce, T.  
TITLE Direct Submission  
SUBMITTED (01-APR-1993) Toshiki Noce, Mitsubishi Kasei Institute of Life Sciences, Developmental Biology, 11 Minamiooya, Machida, Tokyo 194, Japan (Tel:0427-24-6246, Fax:0427-29-1252)  
LOCATION/Qualifiers  
1. 1930  
/organism="Mus musculus"  
/strain="BALB/c"  
/sub\_species="domesticus"  
/db\_xref="taxon:10090"  
/cell\_type="primordial germ cell, spermatogonium, spermatocyte"  
/tissue\_type="gonad, testis"  
/clone\_lib="lambda gt10"  
/dev\_stage="adult, embryo"  
1. 1914  
/function="RNA helicase"  
/codon\_start=1  
/product="Drosophila vasa homologue"  
/protein\_id="BA03584.1"  
/db\_xref="GI:286075"  
/translation="GGRGKGFNGRFLNKEEGDSSGFWKESNNDCEDNQRSGFS KRGGCDNDSEASGPRGGRGFRGCGFGLRPSNEDDQGTCCGGFVLGK PAASDNGDYSRSGSGRGYKGLNEVYSGSKNWKSETEGESDSSGQPVY IPPPEDEDSIFAHYOTGINDPKYDTLIVESGHDAPALITPEANLCOTLNNNR KAGTYKLTVPVQKTYTPIYLRBDLMACQOTSGKTAFLILILAHMRDGTASRFX LQEPICUVAIPRELINQIYLEARKSFSTCVISVITGGFSGSVQIVQCNILC ATPGRIMDIIGKEKIGLKQVYLVLEADSMIDMFABEIKKLISCPMPKEQHQL LFSATFPEIIRIAGDFLKSNTLVFAVAQVCGACAGAGTCTGGGAGACATGC 1019 NIGDRFWFVETKKADFIATPLQEKISFSIHGDEOREBQALDFFGCKPLFYE VATSVAARGLDIENYOHVINDIPSTIDEVHRTGRCGRCNTGPAISFPTDDNHL AOPLVKUSDAGQDVPANLEIATSTYPPFSSTSGCAVPAVDTRKNYQGAHYZ YSGDFFTSSSSS"  
BASE COUNT 577 a 356 c 487 g 510 t  
ORIGIN  
Query Match 2.3%; Score 50; DB 10; Length 1930;  
Best Local Similarity 100.0%; Pred. No. 1.3e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS AC124894 108795 bp DNA linear HTG 03-JUN-2002  
DEFINITION Rattus norvegicus clone CH230-5A4, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
ACCESSION AC124894  
VERSION AC124894.1 GI:21465340  
KEYWORDS HTG; HTG\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
REFERENCE 1 (bases 1 to 108795)  
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Albrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T., Barbata, J., Benton, J., Bimaga, K., Blackenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, C.D., Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louesged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Matlindate, A., Martinez, E., Massey, E., Mawney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okumou, G., Oragunye, N., Oyale, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojudo, I., Rolle, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoochitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tameria, A., Tameria, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umanai, K., Vasquez, L., Vera, V., Villalona, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinscock, G. and Gibbs, R.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 108795)  
AUTHORS Worley, K.C.  
TITLE Direct Submission  
SUBMITTED (19-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
JOURNAL  
REFERENCE 3 (bases 1 to 108795)  
AUTHORS Worley, K.C.  
TITLE Direct Submission  
SUBMITTED (03-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
COMMENT  
Center: Baylor College of Medicine  
Genome Center  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: GBAA

```

Center clone name: CH230-5A4
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 56701 bases at least Q40
Consensus quality: 62823 bases at least Q20
Consensus quality: 66969 bases at least Q20
-----
NOTE: Estimated insert size may differ from sequence length
[see http://www.hgsc.bcm.tmc.edu/docs/genbank_drafc_data.html].
NOTE: This is a 'working draft' sequence. It currently
consists of 56 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
1168: contig of 1168 bp in length
1268: gap of unknown length
1269: contig of 1093 bp in length
2352: gap of unknown length
2461: gap of unknown length
3595: contig of 1134 bp in length
3596: gap of unknown length
3696: gap of unknown length
5046: contig of 1351 bp in length
5047: gap of unknown length
5147: contig of 1110 bp in length
6257: gap of unknown length
6356: gap of unknown length
7858: contig of 1502 bp in length
7859: gap of unknown length
7959: gap of unknown length
9544: contig of 1586 bp in length
9545: gap of unknown length
9645: contig of 1204 bp in length
10849: gap of unknown length
10948: contig of 1807 bp in length
12755: gap of unknown length
12756: contig of 1440 bp in length
14295: gap of unknown length
14296: contig of 1113 bp in length
14396: gap of unknown length
15509: gap of unknown length
15608: contig of 1184 bp in length
16792: gap of unknown length
16793: contig of 1843 bp in length
16893: gap of unknown length
18735: contig of 1351 bp in length
18736: gap of unknown length
20186: contig of 1602 bp in length
20187: gap of unknown length
21888: contig of 1795 bp in length
21889: gap of unknown length
23783: contig of 1727 bp in length
23784: gap of unknown length
25610: contig of 2170 bp in length
25611: gap of unknown length
27880: contig of 1080 bp in length
27881: gap of unknown length
29060: contig of 1426 bp in length
29061: gap of unknown length
30586: contig of 1720 bp in length
30587: gap of unknown length
32406: contig of 1784 bp in length
32407: gap of unknown length
34290: contig of 2313 bp in length
34291: gap of unknown length
36703: contig of 1027 bp in length
36704: gap of unknown length
37930: contig of 1087 bp in length
37931: gap of unknown length
39017: contig of 1020 bp in length
39018: gap of unknown length
40137: contig of 2152 bp in length
40138: gap of unknown length
42389: contig of 2152 bp in length
42390: gap of unknown length

```

```

42490 43589: contig of 1100 bp in length
43590 43689: gap of unknown length
43690 45802: contig of 2113 bp in length
45803 45902: gap of unknown length
45903 48014: contig of 2112 bp in length
48015 48114: gap of unknown length
48115 49732: contig of 1618 bp in length
49733 49832: gap of unknown length
49833 51415: contig of 1583 bp in length
51416 51515: gap of unknown length
51516 52858: contig of 1343 bp in length
52859 52958: gap of unknown length
52959 55075: contig of 2117 bp in length
55076 55175: gap of unknown length
55176 57313: contig of 2137 bp in length
57313 57412: gap of unknown length
57412 59494: contig of 2082 bp in length
59494 59495: gap of unknown length
59495 61146: contig of 1552 bp in length
61146 61247: gap of unknown length
61247 63268: contig of 2022 bp in length
63268 63368: gap of unknown length
63369 65294: contig of 1926 bp in length
65294 65394: gap of unknown length
65395 67458: contig of 2063 bp in length
67458 67557: gap of unknown length
67557 69843: contig of 2286 bp in length
69843 69943: gap of unknown length
69943 72337: contig of 2394 bp in length
72337 72437: gap of unknown length
72437 73987: contig of 1550 bp in length
73987 74087: gap of unknown length
74087 76226: contig of 2139 bp in length
76226 76326: gap of unknown length
76326 79288: contig of 2962 bp in length
79288 79388: gap of unknown length
79388 81977: contig of 2589 bp in length
81977 82077: gap of unknown length
82077 84162: contig of 2085 bp in length
84163 84262: gap of unknown length
84263 87197: contig of 2935 bp in length
87197 87298: gap of unknown length
87298 90154: contig of 2856 bp in length
90154 90253: gap of unknown length
90253 92307: contig of 2054 bp in length
92307 92407: gap of unknown length
92407 94844: contig of 2437 bp in length
94844 94944: gap of unknown length
94945 97858: contig of 2914 bp in length
97859 97959: gap of unknown length
97959 100200: contig of 2242 bp in length
100201 100300: gap of unknown length
100301 102763: contig of 2463 bp in length
102763 102863: gap of unknown length

Query Match      2.2%: Score 47; DB 2; Length 108795;
Best Local Similarity 100.0%; Pred. No. 4; 8e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      757 GAGATGAGACATCTTTCACATTTATCAGACGACGATAACTT 803
Db      89719 GAGATGAGACATCTTTCACATTTATCAGACGACGATAACTT 89673

RESULT 14
AC105884/c      114024 bp      DNA      linear      HTG 13-JUL-2002
LOCUS          Ratius norvegicus clone CH230-1302, *** SEQUENCING IN PROGRESS ***
DEFINITION     52 unordered pieces.
ACCESSION     AC105884
VERSION       AC105884.2   GI:21736908
KEYWORDS       HTG; HTGS PHASE1.
SOURCE         Norway rat.

```

## ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE  
AUTHORS

1 (bases 1 to 114024)  
Munry,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
David,M.L., Davis,C., Davy-Carroll,L., Dederich,D.H.,  
DeJany,K.R., Delgado,O., Dem,A.L., Ding,Y., Dink,H.,  
Douthaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,  
Gabris,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Hosni,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leat,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtege,O., Lieu,C., Liu,J., Liu,W., Loussaged,H.,  
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogun,M., Okunolu,G.,  
Ozgunye,N., Oyedero,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Petere,L., Pickens,R., Primus,E., Pu,L.B., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojoubokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scheier,S., Scott,G., Shen,H., Shoochari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umanil,K., Vasquez,U., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

TITLE  
JOURNAL

Unpublished  
2 (bases 1 to 114024)

REFERENCE  
AUTHORS  
JOURNAL

Submitted (10-JAN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18104791.

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GNZK  
Center clone name: CH230-11302  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 62592 bases at least Q40

Consensus quality: 66494 bases at least Q30  
Consensus quality: 69970 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length.  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 52 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1	1255:	contig of 1255 bp in length
*	1256:	gap of unknown length
*	1356:	contig of 1557 bp in length
*	2913:	gap of unknown length
*	3013:	contig of 1486 bp in length
*	4499:	gap of unknown length
*	4599:	contig of 1291 bp in length
*	5889:	gap of unknown length
*	5990:	contig of 1589 bp in length
*	7579:	gap of unknown length
*	7679:	contig of 1348 bp in length
*	9026:	contig of 1060 bp in length
*	9126:	gap of unknown length
*	9127:	contig of 1060 bp in length
*	10187:	gap of unknown length
*	10287:	contig of 1087 bp in length
*	11374:	gap of unknown length
*	11474:	contig of 1556 bp in length
*	13030:	gap of unknown length
*	13130:	contig of 1538 bp in length
*	14668:	gap of unknown length
*	14768:	contig of 1468 bp in length
*	16236:	gap of unknown length
*	16335:	contig of 1022 bp in length
*	16336:	gap of unknown length
*	17357:	contig of 1402 bp in length
*	17458:	gap of unknown length
*	18859:	contig of 1719 bp in length
*	18960:	gap of unknown length
*	20779:	contig of 1146 bp in length
*	20779:	gap of unknown length
*	21924:	contig of 1032 bp in length
*	22025:	gap of unknown length
*	22025:	contig of 1195 bp in length
*	23057:	gap of unknown length
*	23157:	contig of 1195 bp in length
*	24352:	gap of unknown length
*	24452:	contig of 1211 bp in length
*	25663:	gap of unknown length
*	25763:	contig of 1406 bp in length
*	27169:	gap of unknown length
*	27268:	contig of 1107 bp in length
*	28375:	gap of unknown length
*	28475:	contig of 1761 bp in length
*	28476:	gap of unknown length
*	30236:	contig of 1324 bp in length
*	30337:	gap of unknown length
*	31661:	contig of 1347 bp in length
*	31760:	gap of unknown length
*	33107:	contig of 1347 bp in length
*	33108:	gap of unknown length
*	33207:	contig of 2618 bp in length
*	33208:	gap of unknown length
*	35825:	contig of 1479 bp in length
*	35925:	gap of unknown length
*	35926:	contig of 1479 bp in length
*	37404:	gap of unknown length
*	37504:	contig of 1765 bp in length
*	37505:	gap of unknown length
*	39270:	contig of 2165 bp in length
*	39369:	gap of unknown length
*	41534:	contig of 1428 bp in length
*	41634:	gap of unknown length
*	43062:	contig of 2056 bp in length
*	43063:	gap of unknown length
*	43163:	contig of 1571 bp in length
*	45218:	gap of unknown length
*	45318:	contig of 1571 bp in length
*	45319:	gap of unknown length
*	46889:	contig of 1571 bp in length
*	46890:	gap of unknown length



```

* 46990 48117: contig of 1128 bp in length
* 48118 48217: gap of unknown length
* 48218 50706: contig of 2489 bp in length
* 50707 50806: gap of unknown length
* 50807 52398: contig of 1592 bp in length
* 52399 52498: gap of unknown length
* 52499 54416: contig of 1918 bp in length
* 54417 54517: gap of unknown length
* 54517 56139: contig of 1623 bp in length
* 56140 56239: gap of unknown length
* 56240 58833: contig of 2594 bp in length
* 58834 58934: gap of unknown length
* 58934 61254: contig of 2321 bp in length
* 61255 61354: gap of unknown length
* 61355 63634: contig of 2280 bp in length
* 63635 63734: gap of unknown length
* 63735 66620: contig of 2886 bp in length
* 66621 66720: gap of unknown length
* 66721 68735: contig of 3015 bp in length
* 68736 69835: gap of unknown length
* 69836 73022: contig of 3186 bp in length
* 73022 73121: gap of unknown length
* 73121 75151: contig of 2030 bp in length
* 75152 75252: gap of unknown length
* 75252 77299: contig of 2048 bp in length
* 77300 77399: gap of unknown length
* 77400 80847: contig of 3448 bp in length
* 80848 80947: gap of unknown length
* 80948 82958: contig of 2011 bp in length
* 82959 83058: gap of unknown length
* 83059 86408: contig of 3350 bp in length
* 86409 86508: gap of unknown length
* 86509 89930: contig of 3422 bp in length
* 89931 90030: gap of unknown length
* 90031 95181: contig of 5151 bp in length
* 95182 95281: gap of unknown length
* 95282 100600: contig of 4779 bp in length
* 100601 100160: gap of unknown length
* 100161 103757: contig of 3596 bp in length
* 103757 103856: gap of unknown length
* 103857 108054: contig of 4198 bp in length
* 108055 108155: gap of unknown length
* 108155 114024: contig of 5870 bp in length.
Location/Qualifiers
1. 114024
/organism="Rattus norvegicus"
/db_xref="taxon:10116"

```

```

Query Match 2.2%; Score 47; DB 2; Length 114024;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 757 GAGGATAGGACCTCATTGTCACATTATGACAGCATTAACCTT 803
Db 10537 GAGGATAGGACCTCATTGTCACATTATGACAGCATTAACCTT 10491

```

```

RESULT 15
AE013856/c 14350 bp DNA linear BCT 26-JUL-2002
DEFINITION Yersinia pestis KIM section 256 of 415 of the complete genome.
ACCESSION AE013856 AE009992
VERSION AE013856.1 GI:21959401
KEYWORDS
SOURCE
ORGANISM
Yersinia pestis KIM.
Yersinia pestis KIM.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Yersinia.

```

```

REFERENCE
AUTHORS
1 (bases 1 to 14350)
Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F.,
Lies P., Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
Fetherston J.D., Lindler L.E., Brubaker R.R., Plana G.V.,
Straley S.C., McDonough K.A., Nilles M.L., Matson J.S.,

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F.,
Lies P., Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
Fetherston J.D., Lindler L.E., Brubaker R.R., Plana G.V.,
Straley S.C., McDonough K.A., Nilles M.L., Matson J.S.,
Blattner F.R. and Perry R.D.
Direct Submission
Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. 114350
/organism="Yersinia pestis KIM"
/strain="KIM"
/db_xref="taxon:187410"
complement(78..1460)
/gene="dbpA"
/note="Y2531"
complement(78..1460)
/gene="dbpA"
/function="enzyme; RNA synthesis, modification, DNA
transcription"
/note="residues 4 to 460 of 460 are 68.84 pct identical to
residues 2 to 457 of 457 from E. coli K12 : B143;
residues 4 to 460 of 460 are 68.66 pct identical to
residues 2 to 457 of 457 from GenPept :
>gi|AG56454.1|AE005371.10 (AE005371) ATP-dependent RNA
helicase [Escherichia coli O157:H7 EDL933]"
/codon_start=1
/transl_table=1
/product="ATP-dependent RNA helicase"
/protein_id="AA086087.1"
/db_xref="GI:21959402"
/translation="MSTTSPSSLTLPAAQSLNINELGYTENTPQAAVLPAALNGDY
RAKAKTSGTAAFGILGIDKIAVGEFTQALVCPREIADQVSKELRIARFTQNI
KILTCGSGPMGQGLDSLVAPHVVGTPGRIQHLKKTLLDELTLVDEDRML
DMGSDAIDIVIAVTPROTTLFSAVPSGIGTISARVOROPINAGVGDDEPAIE
QVETTRERKRLPILISLVHYOPASCVVCNTRKDCQSYSESIGISVLAIGDL
EQRDDVLAIFPARRSCRLVATDVARGDIDLELVINFELEFDEIHLIRGRTG
RAGSGAVSLCTPQEWRAHTIEDYQIKLKTMPARQVSRSTIMLEPEVNTLCIG
GRKAKIRPDILGALTGDAGITADVGKIDMFVHAYVAIRKASAKRALQOLQGKIK
GXNCKARLK"
1773..2954
/gene="purT"
/note="Y2533"
1773..2954
/gene="purT"
/function="enzyme; purine ribonucleotide biosynthesis"
/note="residues 1 to 392 of 393 are 78.31 pct identical to
residues 1 to 392 of 392 from E. coli K12 : B1849"
/codon_start=1
/transl_table=1
/product="phosphoribosylglycinamide formyltransferase 2"
/protein_id="AA086088.1"
/db_xref="GI:21959403"
/translation="MLTGTALRPGATRWMLGAGELKEVAIECQRLGLEVAVDRY
ADAPAMVAVRSHVINMLDGAALQVAAQERPHYIVETIAIADMLVELEKQOHV
PCAETRLTNRRIGIRLAEITLQLPSSYRFDVDSAFQAVDIDYPCIVKPMVS
SGKQGLIRSEEHLOAMEYAAQCGRAGSRVIEGVHFDFTLTLIRAVDSIHFC
APIGRQEDDYRESNQPMASDIALORAKETISAOVYTAGRGELFGVELFVCGDDVI
FSEVSRPHDTGVTTLISQMSERLHVRAFLGIPITIOYGAASAVILPELTSON
ITRGLETALIGDTQIRLFGKPEIAGRRRLGVALAVADNLETALVAKKAAGNTEVSG
E"
complement(2151..2369)
/gene="Y2532"
complement(2151..2369)
/gene="Y2532"
/function="unknown"
/note="residues 41 to 68 of 72 are 46.42 pct identical to
residues 841 to 868 of 1147 from GenPept :

```

```

gene
CDS

```

```

>emb|CAB40374.1| (AJ225088) Starch synthase isoform SS III
/vigna unguiculata) "
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AA086089.1"
/db_xref="GI:21959404"
/translation="MNOTFDNADPATGATATLTCIFPSLKVFLAADAQALAFTRRTHYR
FNARVTVVAGLKEGALCVGKAVTGKG"
/complement(3080..3337)
/gene="Y2534"
/complement(3080..3337)
/gene="Y2534"
/notes="residues 1 to 59 of 85 are 59.32 pct identical to
residues 1 to 58 of 59 from E. coli K12 : B1811; residues
1 to 59 of 85 are 62.71 pct identical to residues 1 to 58
of 59 from Genpept : >gb|AA020738.1| (AB008781) putative
cytoplasmic protein [Salmonella typhimurium LT2]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA086090.1"
/db_xref="GI:21959405"
/translation="MLAGMPLSHBEOQEAVERJHKFMSGMSGGAIALVAAEIRER
HONDPQMAIFEDHDFDEHTSDYRDDEPDADIEDPYEG"
/3502..4878
/gene="pabB"
/notes="Y2535"
/3502..4878
/gene="pabB"
/function="enzyme; biosynthesis of cofactors, carriers:
Folic acid"
/notes="residues 11 to 452 of 458 are 66.81 pct identical
to residues 11 to 451 of 453 from E. coli K12 : B1812;
residues 7 to 452 of 458 are 67.33 pct identical to
residues 8 to 452 of 454 from Genpept : >gb|AA020739.1|
(Salmonella typhimurium LT2)"
/codon_start=1
/transl_table=11
/product="p-aminobenzoate synthetase, component I"
/protein_id="AA086091.1"
/db_xref="GI:21959406"
/translation="MCEKELFYKELPYHPDALLOOFTPLANQAMMLHSGFAQPHS
RFDILVADPKITITHTGDKTIVSDQGETSLADPFSLLQOOLKRPAPAMADHLP
OGGALGLFGYLDGRVVEKLPFLAANDIALPMAICGLYWMALIDHSRQKLTIVNSGA
EORLHWLHOQTIDENLVLPKULVQWQUMSRSEQGEKRFQIQAIRSDCYQIRLAKR
FSAEYQDEMOAFSLSRNQAPFSAFRLPDNALISVPERFLMENVHVOYQTRIKG
TLPRADPEODROQAERLANSAKQDAENLMTVDLRNDIGVARGSVPELFEVPEP
FPVAVHLVSTITAILPEPCSPTELLRACFGSGITGAKRVAMEIIEOLEPVRNAYC
GNIGYISCGGMDNITIRITLMTENGKICYCAGGIVADSQEQAEVCEYCFPKVARILP
OLGECVIS"
/4875..5474
/gene="Y2536"
/4875..5474
/gene="Y2536"
/notes="residues 10 to 199 of 199 are 55.78 pct identical
to residues 8 to 190 of 192 from E. coli K12 : B1813;
residues 10 to 199 of 199 are 55.26 pct identical to
residues 8 to 190 of 192 from Genpept : >gb|AA020740.1|
(AE008781) putative NTP pyrophosphohydrolase [Salmonella
typhimurium LT2]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA086092.1"
/db_xref="GI:21959407"
/translation="MSELITGOYLSEFINRFOLOPPDNVLTSHYSFATNRAAVL
IPIICRPPTLITRRADHTRKAGQVAPPGKADPDOSLISTALREAEVATIPAS
VHVVLGKLAPLNSSGCHVTPIVGCVPANIPYGNDEYAGLFEIPLHBAISLSKYHS
LDIHREGINHRYLTSWENQITWGLTATITIRHLAQOVS"
/6182..7546
/gene="sdaA"

```

```

CDS
/notes="Y2537"
/6182..7546
/gene="sdaA"
/function="enzyme; degradation of small molecules; amino
acids"
/notes="residues 1 to 453 of 454 are 84.32 pct identical to
residues 1 to 453 of 454 from E. coli K12 : B1814"
/codon_start=1
/transl_table=11
/product="L-serine deaminase"
/protein_id="AA086093.1"
/db_xref="GI:21959408"
/translation="MISVDFMFKIGIGPSSSHTVGPKAKQGFVDLLITGLMPSITR
VAIDVYGSLSITGKGHTDIAIMGLAGNLPDVIDIDSLPAFIRVELRQKMLANGL
HEVDVPEGGWVRSDNLPLHENGMOIHAQGEDEKSTKYSIGGFIIDENFGKA
SVNDVAVPYFNSAAELIANVCEOTGISGMWONELAMSKESIEISFYAIQOTMRA
CIRGLNTEBVLGSPRVRPRAALRLIYSSDKLSDDPIVIVDVMFPLAINEEA
AGRVVATPNTGACGIVPAVLAYDHTFEVTEIFRIRPLASGALICILKMAASISG
AEVCGOEVGACSMMAAGLAEILGASPIQVCIAGTAEIGMEHNLGLTCDPVAGOVPC
IERNALASVKAINSARMAMRTSEPRVSLDKVLETMETEKDMNAKYRETSRGLAIK
VOCT"
/complement(7826..8323)
/gene="Y2538"
/complement(7826..8323)
/notes="residues 15 to 164 of 165 are 47.33 pct identical
to residues 1 to 148 of 152 from E. coli K12 : B1007;
residues 1 to 165 of 165 are 73.93 pct identical to
residues 1 to 165 of 172 from Genpept : >gb|AA008740.1|
(AF021839) 4-hydroxyphenylacetic acid hydroxylase putative
coupling protein [Photobacterium luminescens]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA086094.1"
/db_xref="GI:21959409"
/translation="MSTENHHCFRDMACLSAANVITVTDEPAGCGTITATVCSV
TDIPTPLIICINRNSANPVQONGCICIVLNHQDPMARHFGMGSSMEERFSWD
IMNVGLGQPLRLDRLTSLLEGVIEOVQITGTHLVYVITQIVLREGHGILYFKRHF
HPVWV"

```

Query Match 1.7%; Score 36; DB 1; Length 14350;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-08;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1333 TTGGATGAAGCTGATCGCATGTTGATATGGCTTT 1368
Db 1004 TTGGATGAAGCTGATCGCATGTTGATATGGCTTT 969

```

Search completed: June 10, 2003, 14:05:22  
 Job time : 5529 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:06:30 ; Search time 87 seconds  
(without alignments)  
10288.146 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 3954  
Sequence: 1 atggggagatgaagatgtgga.....tagatgatgatcatggat 2172

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 671580 segs, 206047115 residues  
Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p.model -DEV=xip  
-O/cgmr2.1/usfpo.spool/US09714865/runat\_05062003.111758\_26020/app.query.fasta.1.2311  
-DB=SPREMBL\_21 -QMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=Disum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09714865 @CGN 1.1.119 @runat 05062003.111758\_26020 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTPUT -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DET\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_rvitus:\*  
17: sp\_bacteriaph:\*  
18: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	2001.5	50.6	715 13 Q8QG8	Q8QG8 brachydanio

2	1997	50.5	700 13	Q42378	Q42378 brachydanio
3	1997	50.5	716 13	Q42107	Q42107 brachydanio
4	1995.5	50.5	715 13	Q8QF00	Q8QF00 brachydanio
5	1959.5	49.6	700 13	Q91372	Q91372 xenopus lae
6	1899.5	48.0	617 13	Q902F6	Q902F6 cyrtias lat
7	1885	47.7	647 13	Q9PT10	Q9PT10 oncorhynch
8	1873	47.4	621 13	Q8QHL6	Q8QHL6 oreochromis
9	1865	47.2	645 13	Q9DEG3	Q9DEG3 oreochromis
10	1763.5	44.6	662 13	Q9DGR8	Q9DGR8 gallus gall
11	1570	39.7	770 5	Q9GNP1	Q9GNP1 ciona savig
12	1563	39.5	688 5	Q9GNP2	Q9GNP2 ciona savig
13	1522	38.5	659 5	Q96069	Q96069 ciona lites
14	1522	38.5	669 5	Q96068	Q96068 ciona lites
15	1497	37.9	394 13	Q8QGD0	Q8QGD0 pantodon bu
16	1489.5	37.7	399 13	Q918L8	Q918L8 danio dangi
17	1473.5	37.3	397 13	Q8QGD2	Q8QGD2 hyphessobry
18	1460	36.9	388 13	Q8QGD3	Q8QGD3 cyrtinus ca
19	1454	36.8	336 13	Q8QGC8	Q8QGC8 oncorhynch
20	1447	36.6	400 13	Q8QGC9	Q8QGC9 melanotaeni
21	1425.5	36.1	546 5	Q9GV10	Q9GV10 ephydatia f
22	1412	35.7	396 13	Q8QGD1	Q8QGD1 cryzias lat
23	1388	35.1	797 5	Q9GV13	Q9GV13 hydra magni
24	1277	32.3	601 5	Q01378	Q01378 bombyx mori
25	1257.5	31.8	830 5	Q9GV12	Q9GV12 hydra magni
26	1242.5	31.4	798 5	Q9VHP0	Q9VHP0 drosophila
27	1239.5	31.3	798 5	Q8SX18	Q8SX18 drosophila
28	1238	31.3	688 13	Q42375	Q42375 brachydanio
29	1202	30.4	658 11	Q9QWS9	Q9QWS9 mus musculus
30	1179.5	29.8	628 5	Q9GV14	Q9GV14 hydra magni
31	1176	29.7	688 3	Q9PE09	Q9PE09 neurospora
32	1165	29.5	617 3	Q8TFK8	Q8TFK8 candida gla
33	1147.5	29.0	726 5	Q97031	Q97031 dugesia jap
34	1128.5	28.5	641 5	Q9N3F4	Q9N3F4 caenorhabdi
35	1127	28.5	644 5	Q9N3Y1	Q9N3Y1 caenorhabdi
36	1100	27.8	633 10	Q9SIN6	Q9SIN6 arabidopsis
37	1095.5	27.7	646 10	Q9M2F9	Q9M2F9 arabidopsis
38	1094.5	27.7	763 5	Q22873	Q22873 caenorhabdi
39	1093	27.6	781 5	Q97032	Q97032 dugesia jap
40	1075	27.2	974 5	Q96619	Q96619 caenorhabdi
41	1075	27.2	974 5	Q27376	Q27376 caenorhabdi
42	1069.5	27.0	603 10	Q9M2G5	Q9M2G5 arabidopsis
43	1066.5	27.0	491 5	Q9GV11	Q9GV11 ephydatia f
44	1017.5	25.7	573 5	Q9GV07	Q9GV07 dugesia dor
45	999	25.3	376 5	Q8SXU8	Q8SXU8 drosophila

## ALIGNMENTS

RESULT 1  
ID Q8QG8 PRELIMINARY; PRT; 715 AA.  
AC Q8QG8;  
DT 01-JUN-2002 (TEMBLrel. 21, Created)  
DT 01-JUN-2002 (TEMBLrel. 21, last sequence update)  
DT 01-JUN-2002 (TEMBLrel. 21, last annotation update)  
DE Vasa-like protein.  
GN VASA.  
OS Brachydanio rerio (zebrafish) (zebra danio).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB;  
RA Bartal R., Orban L.,  
RT "Characterization of the genomic locus encoding vasa protein in zebrafish."  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF461759; AAL89410.1; -  
SQ SEQUENCE 715 AA; 76795 MW; 602331B3700203AA CRC64;

Alignment Scores:



```

Db      638 ThrProLeuAlaArgSerLeuValValLeuSerGlyAlaGlnGlnValValProLys 657
QY      2008 TGGTGGAGAAATTCGCTTAGACATACATTCCTCGCTTCAGTGT---AGTACAGA 2064
Db      658 TrpLeuGlnValAlaAlaPheSerAlaHis-----GlyThrThrGlyPheAsnProArg 675
QY      2065 GGAACGCTGTTGTCATGATTGATACAGAAAGGCAAGACACTTTGAACACAGCTGG 2124
Db      676 GlyValValPheAlaSerThrAspSerArgLysGlyGlySer-----AATCCAGTA 2154
QY      2125 TTTTCTCTTCACAGACTCC-----
Db      690 PheLysSerAspGluProProPheSerGlnThrSerAlaProSerAlaAlaAlaAla 709
QY      2155 GATGATGATCATGGGAT 2172
Db      710 AspAspGlnGluTrpGln 715

RESULT 2
ID      042378      PRELIMINARY;      PRT;      700 AA.
AC      042378;
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE      RNA helicase (DEAD box).
GN      VASA OR VLG.
OS      Brachydanio rerio (Zebrafish) (Zebra danio).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_Taxid=7955;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98025484; PubMed=9376327;
RA      Olsen L.C., Aasland R., Elose A.;
RT      "A vasa-like gene in Zebrafish identifies putative primordial germ
RT      cells.";
RL      Mech. Dev. 66:95-105 (1997).
CC      1.- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR      EMBL; Y11007; CAAT7273.1; -.
DR      HSSP; Q58083; 1HV8.
DR      ZFIN; ZDB-GENE-990415-272; vasa.
DR      InterPro; IPR001410; DEAD.
DR      InterPro; IPR000629; DEAD box.
DR      InterPro; IPR001650; Helicase_C.
DR      Pfam; PF00270; DEAD_1.
DR      Pfam; PF00271; Helicase_C_1.
DR      SMART; SM00487; DEXDC; 1.
DR      SMART; SM00490; HELICG; 1.
DR      PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW      ATP-binding; Helicase; RNA-binding.
SQ      SEQUENCE 700 AA; 75312 MW; 726B56A90DC62DB3 CRC64;

Alignment Scores:
Pred. No.:      2,34e-147      Length:      700
Score:      1997.00      Matches:      423
Percent Similarity:      67.37%      Conservative:      89
Best Local Similarity:      55.66%      Mismatches:      148
Query Match:      50.51%      Indels:      100
DB:      13      Gaps:      20

US-09-714-865-15 (1-2172) x 042378 (1-700)
QY      10 GAAGATTGGAGAGCAAAATCAACCTTCATATGCTTCCTCATATGTTCCCATATTGAGAG 69
Db      2 AspAspTrpGlnGluAspGlnSerProValValSerCys-----Ser 15
QY      70 GATGAGATTCTGGAGAAATGAGACATATTTAAACAGAGACTCCAGCTTCATCATGAA 129
Db      16 SerGlyPheGlyGlyAlaGlyAsnAspLys-----SerAsnSerGlu 29

```

```

QY      130 ATGATGATGACCTTCTCGAAGATCATTTGATGAAAGTGATTTGCTCTGGCCG 189
Db      30 GlyThrGlnLysSerSerTrpLys-----MetThrGlyAsp 41
QY      190 AATTGGAACACGAGATGCTGGTGAAGTAAATAGCGAGATATATCATCCAAATGGT 249
Db      42 SerPheArgGlyArg-----Gly 47
QY      250 GATTTCAGAGTTGGAAGAGTTTGGAAACAGAGTTTCA-----AACAGCAGGTTT 303
Db      48 GlyArgGly---GlySerArgGlyGlyArgGlyGlyPheSerGlyPheLysSerGlnIle 66
QY      304 GAAGATGATGATAGCTCTGCTTTCTGAGAGAGCTAGTAATGACCTGCGAATATATCA 363
Db      67 AspGlnAsnGlySerAspGlyGlyTrp-----AsnGlyGlyGlnSerArgGly 82
QY      364 ACAAGCAACAGAGGGTTTTCAGAGAGCGGCTATCGAGATGGAATTAATTCAGAA--- 420
Db      83 ArgGlyArgGlyGlyPhe-----ArgGlyGlyPheArgSerGlySerArgAspGluAsn 100
QY      421 -----GCTTCAGGGCCATACAGAGAGGTGAGAGGT 453
Db      101 AspGlnAsnGlyAsnAspAspGlyTrpLysGlyGlyGlnSerArgGlyArgGly 120
QY      454 AGTTTC---CGAGTTGCCCTGAGAGATT----- 480
Db      121 GlyPheGlyGlyGlyPheArgGlyGlyPheArgAspGlyGlyAsnGluAspThrGlyArg 140
QY      481 ---GGTTCAGGAAGTCCAAATATGACTTAGACCCAGACGAATGTATGACG----- 528
Db      141 ArgGlyPheGlyArgGlnAsnGlnAsnGlyAsnAspGlnGlyGlyArgGly 160
QY      529 ---CGACCTGGGCGCTTTGGTTCTAGAGACCAAGTATTAAGTGCCACAGGTAATGT 585
Db      161 ArgGlyArgGlyGlyPheArgGlyGlyPheArg-----AspGlyGlyGly 175
QY      586 GATACCTTCTCAAGCAAGAGTGCAGTGAAGTGAACGAGTGTTACAAAGTTTAAT 645
Db      176 AspGlnSerGlyLysArg-----GlyPheGlyArgGlyGlyPheArgGlyArgAsn 192
QY      646 GAAGAGTAATTAACAGGCTCTGGAAGAAATTTCTTGAAGTCAAGACAGAGAGAGAA 705
Db      193 GlnGlnValPheSer---LysValThrThrAlaAspLysLeuAspGlnGlnGlySerGln 211
QY      706 AGTAGATATCTCAAGAACCAAGTACCTACATACCCCTCCACCTGAGAGAG 765
Db      212 AsnAla-----GlyProLysValValTrpValProProProProGlnGln 228
QY      766 GACTCCATCTTTCACATTAATACAGACAGCATTAACCTTCAGCAAAATTCACACTATTCT 825
Db      229 SerSerIlePheSerHisTrpAlaThrGlyIleAsnPheAspLysTrpAspAlaIleLeu 248
QY      826 GTGAGAGTGTCTGACATGATGACACCAACGACAAATCTGACTTTGAAGAGCTAATCTC 885
Db      249 ValAspValSerGlySerAsnProProLysAlaIleMetThrPheGlnGlnAlaGlyLeu 268
QY      886 TGTTCAGACATGAATAACACATTCGTAAGCTGGTATTAAGCTTACTCTGTGCAA 945
Db      269 CysAspSerLeuSerLysAsnValSerLysSerGlyTrpValLysProThrProValGln 288
QY      946 AAATACAGATTCTTCATCATCTTGCAGAGAGAGATTGATGAGCTTGCTCAACAGAG 1005
Db      289 LysHisGlyIleProIleIleSerAlaGlyArgAspLeuMetAlaCysAlaGlnThrGly 308
QY      1006 TCTGGGAAGCTCGCGCTTTTCTCTCAACAAATTTGGCTCATATGATGATGAGATA 1065
Db      309 SerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnArgPheMetThrAspGlyVal 328
QY      1066 ACTGACAGCTGTTTAAAGAGCTTGAGAGACCAAGAGTATTTATGAGACCAACCTGCA 1125
Db      329 AlaAlaSerLysPheSerGlnIleGlnGlnProGlnAlaIleIleValAlaProThrArg 348
QY      1126 GAATGTGTCAACAGATTATTTGGAAGCCAGAAATTTCTTTTGGACTGTGTAGA 1185

```

```

Db      349 GluLeuIleAsnGlnIleTyrLeuGluIuAlaArgYsPheAlaTyrGlyThrCysValArg 368
Qy      1186 GCTGTGTATATATGGGGGACCAGCTGGAGATTCATTCACAAATATGTCAAGGC 1245
Db      369 ProValValValTyrGlyGlyIleAsnThrGlyThrThrIleArgGluValLeuValGly 388
Qy      1246 TGTATATATATATGCTACTCTGTAAGACTGATGATGATCATGAGCAAAAGAAATTT 1305
Db      389 CysAsnValLeuCysAlaThrProGlyArgLeuHisAspLeuIleGlyArgGlyValIle 408
Qy      1306 GGTCTCAACAGATCAAAATACTAGTTTGGATGAGCTGATCCGATGTTGATATGAGT 1365
Db      409 GlyLeuSerLysValArgTyrLeuValLeuAspGluIuAlaAspArgMetLeuAspMetGly 428
Qy      1366 TTGTGTCACAAATGAAGAACTTAATTTCTTCCAGAGATGCCATCAAGAGAACACGCC 1425
Db      429 PheGluProGluMetArgLysLeuValAlaSerProGluMetProSerLysGluLysArg 448
Qy      1426 CAAACCTTATGTCAGTCAACTTTTCCAGAGAAATCAAGAGATGCTGAGAGTTT 1485
Db      449 GlnThrLeuMetPheSerAlaThrTyrProGluAspIleGlnArgMetAlaAlaAspPhe 468
Qy      1486 TTAAGTCAAATTAATCTGTTTGTGCTGTTGACAACTGGGTGAGCATGTAGATGTT 1545
Db      469 LeuLysValAspTyrIlePheLeuAlaValGlyValValGlyGlyAlaCysSerAspVal 488
Qy      1546 CAGACAGCCCTTCCCAAGTTGGCCAGTCTCTCAAAAAGAGAAAAGCTGTTGAATCTG 1605
Db      489 GlnGlnThrIleValGlnValAspGlnTyrSerLysArgAspGlnLeuLeuLeu 508
Qy      1606 CGAAATCAGAGGAGTGAAGAAGCATGCTGTTGTAAGTAAGAAAAGACAGATTTT 1665
Db      509 ArgAlaThrGlyAsnGluArgThrMetValPheValGluThrLysArgSerAlaAspPhe 528
Qy      1666 ACTGCAACTTTCTTGTCAAGAAAATATCAATCAAGATTCATGCTGATGGGAA 1725
Db      529 IleAlaThrPheLeuCysGlnGluLysIleSerThrThrSerIleHisGlyAspArgGlu 548
Qy      1726 CAGAGAGCGCGGAGCAAGCTCTTGAGATTTTGGCTTGAAGAGTCCCGCTGTTGTT 1785
Db      549 GlnArgGluArgGluLysValAlaLeuSerAspPheArgLeuGlnHisCysProValLeuVal 568
Qy      1786 GCTACTCAGTAGCTGCAGAGGCTGATATGTAATGTAATGCAACATGTTATCAATTTT 1845
Db      569 AlaThrSerValAlaAlaArgGlyLeuAspIleGluGlnValGlnHisValValAsnPhe 588
Qy      1846 GATCTTCCTTACCATGATGAATATGTCATGCAATGGCGTACTGTCGTTGTGG 1905
Db      589 AspMetProSerSerIleAspGluTyrValHisArgIleGlyArgThrGlyArgCysGly 608
Qy      1906 AATACTGCAGACGATTCCTTTTGTGATCTTGAATGATTAACATTAAGCAGAGCT 1965
Db      609 AsnThrGlyArgAlaValAlaSerPhePheAsnProGluSerAspThrProLeuAlaArgSer 628
Qy      1966 CTAGTAAAGATATGACAGATGCTCAACAGATGTTCTCGCATGTTGAGAGAAATGCC 2025
Db      629 LeuValLysValLeuSerGlyAlaGlnGlnValValProLysTrpLeuGluGluValAla 648
Qy      2026 TTATGATACATTCCTGCTGCTTCAAGTGT--AGTACAGAGAGAAAGCTTTGATCA 2082
Db      649 PheSerAlaHis-----GlyThrThrGlyPheAsnProArgGlyLysValPheAlaSer 666
Qy      2083 GTTGATACCAAGAAAGGCAAGACACTTGAACACAGCTGGCTTTTCTTCCACAGACT 2142
Db      667 ThrAspSerArgLysGlyGlySer-----PheLysSerAspGluPro 680
Qy      2143 CCC-----AATCAGTAGATGAGTCAATGAGAT 2172
Db      681 ProProSerGlnThrSerAlaProSerAlaAlaAlaAlaAlaAspGluGluTyrGlu 700

```

```

ID      042107 PRELIMINARY; PRT; 716 AA.
AC      042107;
DT      01-JUN-1998 (TrEMBLrel. 05, Created)
DT      01-JUN-1998 (TrEMBLrel. 05, Last sequence update)
DE      01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
OS      Vasa.
GN      Brachydanio rerio (Zebrafish) (Zebra danio).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC      Cyprinidae; Danio.
OX      NCBI_TaxID=7955;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97417583; PubMed=9272956;
RA      Yoon C., Kawakami K., Hopkins N.;
RT      "Zebrafish vasa homologue RNA is localized to the cleavage planes of
RT      2-and 4-cell-stage embryos and is expressed in the primordial germ
RT      cells."
RL      Development 124:3157-3166 (1997).
CC      -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
DR      EMBL; AB005147; BAA2253.1; -.
DR      HSP; Q58083; IHW8.
DR      ZFIN; ZDB-GENE-990415-272; vasa.
DR      InterPro; IPR001410; DEAD.
DR      InterPro; IPR000629; DEAD box.
DR      InterPro; IPR001650; Helicase_C.
DR      Pfam; PR00270; DEAD_1; Helicase_C.
DR      Pfam; PF00271; helicase_C_1.
DR      SMART; SM00487; DEXDC; 1.
DR      SMART; SM00490; HELIC; 1.
DR      PROSITE; PS00039; DEAD ATP HELICASE; 1.
KW      ATP-binding; Helicase; RNA-binding.
SQ      SEQUENCE 716 AA; 77004 MW; BB74923B95C7AC5A CRC64;

Alignment Scores:
Pred. No.: 2,35e-147 Length: 716
Score: 1997.00 Matches: 427
Percent Similarity: 66.75% Conservative: 85
Best Local Similarity: 55.67% Mismatches: 157
Query Match: 50.51% Indels: 98
DB: Gaps: 19

US-09-714-865-15 (1-2172) x 042107 (1-716)
Qy      10 GAGATTTGGGAGCAAGAAATCAACCTCATATGTTCTTATGTCATATTTGAGAA 69
Db      2 AspAspTrpGluLysAspGlnSerProValValSer----- 13
Qy      70 GATAGTATCTCGAGAAATGAGACAAATTTTAACAGACATCCAGCTTCATCATCAGAA 129
Db      14 -----CysSerSerGly 17
Qy      130 ATGATGATGACCTTCTCGAAGATCATTTGATGAAAAGTGGATTTGCTTGGCGG 189
Db      18 PheGlyLeuLysSerAsnGlySerAspGlyGlyPheLysSerPheThrGlyGlyAla 37
Qy      190 AATTTTGAACAAGATGCT-----GGTAGTATTAAGCGAGAT 231
Db      38 -----GlyAsnAspLysSerAsnSerGluGlyThrGluLysSerThrTrpLysMetThr 55
Qy      232 AATACATCCCAATGAGCTTTTGGAGTTGGAAGAGTTTGAAGAAAGAAAGAAAGAAAG 288
Db      56 GlyAspSerPheArgGlyArgGlyGlyArgGlyGlySerArgGlyGlyArgGlyPhe 75
Qy      289 TCA-----AACAGCAGTTTGAAGATGATGATGCTGCTGTTTTCGACAGAGCTTACT 342
Db      76 SerGlyPheLysSerGluLysAspGluAsnGlySerAspLysGlyTrp----- 91
Qy      343 AATGACTGCAAGATATATCAACACGGAACAGAGGTTTCCAAAGAGGCGGTATCGA 402
Db      92 AsnGlyGlyLysSerArgGlyArgGlyArgGlyPhe-----ArgGlyGlyPheArg 109

```

QY 403 GATGGAATATTAGAA-----GCTTCAGGCCA 432  
 Db 110 SerGlySerArgAspGluAsnAspGluAsnArgAsnAspAspGlyTyrTyrGlyGlu 129  
 QY 433 TACGAAGAGGTGGAAGGAGTTCGAGGTC---CGTGAGGATTT-----480  
 Db 130 SerArgGlyArgGlyArgGlyGlyPheGlyGlySerPheArgGlyGlyPheArgAspGly 149  
 QY 481 -----GCTTCAGGAATCCAAATTAATGAATTAGACCCAGAC 516  
 Db 150 GlyAsnGluAspThrArgArgArgGlyPheGlyArgGluAsnAsnGluAsnGlyAsnAsp 169  
 QY 517 GAATGATGACG-----CGACTGCTGCTCTTTTGGTTCTAGAACCCAGTA 564  
 Db 170 GluGlyGlyGlyGlyArgGlyArgGlyArgGlyGlyPheArgGlyGlyPheArg-----187  
 QY 565 TTAAAGTGCACAGTAAATGATGATCTTCGAAAGCAAGTGCAGTGCAGAGAACGA 624  
 Db 188 -----AspGlyGlyGlyAspGlySerGlyLysArg-----GlyPheGlyArg 201  
 QY 625 GGTGTTACAAAGGTTTAAATGAAGAATTAACAGGCTCTGAAAGAAATTTCTGAG 684  
 Db 202 GlyGlyPheArgGlyArgAsnGlyGlyValPheSer---LysValThrThrAlaAspLys 220  
 QY 685 TCAGAACGACAGAGAGAGAAAGTAGTATCTCAAGGACCAAAAGTACCTACATACC 744  
 Db 221 LeuAspGlnGlySerGluAsnAla-----GlyProLysValValTyrValPro 237  
 QY 745 CCTCTCCACTGAGAGTGAAGACTCCATCTTTGACATTTTCAGACAGGATTAATCTC 804  
 Db 238 ProProProGluGluGluSerSerLeuPheSerLysTyrAlaThrGlyLysAsnPhe 257  
 QY 805 GACAAATACGACATCTTCTGTGGAAGTCTGTGACATGACACCAACGCAATTCG 864  
 Db 258 AspLysTyrAspAspIleLeuValAspValSerLysSerAsnProProLysAlaIleMet 277  
 QY 865 ACTTTGAAGAGTAACTCTGTCTGACACTGATTAACAACATTTGCTTAAAGCTGTTAT 924  
 Db 278 ThrPheGlnGluAlaGlyLeuCyAspSerLeuSerLysAsnValSerLysSerGlyTyr 297  
 QY 925 ACTAAGCTTACTCTGTGCAAAATACAGTATCTCTATCATATTCGACGACGAGATTG 984  
 Db 298 ValLysProThrProValGlnLysIleGlyIleProIleIleSerAlaGlyArgAspLeu 317  
 QY 985 ATGCTGTGCTCAAAACAGGCTGTGGGAAGCTGGGCTTTCTCCATCAATTTGGCT 1044  
 Db 318 MetAlaCyAsnAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuGln 337  
 QY 1045 CATATGATCATGATGGAATTAAGTCCAGTCTTTTAAAGATTGACAGAACAGAGTGT 1104  
 Db 338 ArgPheMetThrAspGlyValAlaAlaSerLysPheSerGlnIleGlnGluProGluAla 357  
 QY 1105 ATTATTTGACACCAACTCGAATTTGCTCAACAGATTTATTTGGAACCAAAATTT 1164  
 Db 358 IleIleValAlaProThrArgGluLeuIleAsnGlnIleTyrLeuGlnAlaArgLysPhe 377  
 QY 1165 TCTTTGGGACTTGTGTAAGCTGTTGTATATATGAGGAAACCCAGCTGGACATTTCA 1224  
 Db 378 AlaTyrGlnThrCysValArgProValValValTyrGlyGlyLysAsnThrGlyTyrThr 397  
 QY 1225 ATTGCACAAATAGTACAGGCTGTATATATATATGCTACTCTCGAAGACTGATGAT 1284  
 Db 398 IleArgGluValLeuLysGlyCysAsnValLeuCyAsnAlaThrProGlyArgLeuHisAsp 417  
 QY 1285 ATCATAGCAAAAGAAAGATTGCTCTCAACAGATCAAAATCTTATGTTTGGATGAGCT 1344  
 Db 418 LeuIleGlyArgGlyLysIleGlyLeuSerLysValArgTyrLeuValLeuAspGluAla 437  
 QY 1345 GATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404  
 Db 438 AspArgMetLeuAspMetGlyPheGluProGluMetAlaGlyLysLeuValAlaSerProLys 457  
 QY 1405 ATGCCATCAAGAAAGAACAGGCCAAACCTTATGTTCACTGCAATTTTTCAGAGGAATT 1464

Db 458 MetProSerLysGluArgGlnThrLeuMetPheSerAlaThrTyrProGluAspIle 477  
 QY 1465 CAAAGTTGGCTGACAGATTTTAAAGTCAATATATGTTGCTGTTGACAGACTG 1524  
 Db 478 GlnArgMetAlaAlaAspPheLeuLysValAspTyrIlePheLeuAlaValValVal 497  
 QY 1525 GGTGACCATGTAGATGTTTCAGACAGACCGTCTCCAACTGGCCAGTCTCAAAAAGA 1584  
 Db 498 GlyGlyValaCysSerAspValGluGlnThrValValGlnValAspGlnTyrSerLysArg 517  
 QY 1585 GAAAGCTCGTGAATTTCTGCGCAACATAGGAGTAAAGAACTATAGCTTTGTTGAA 1644  
 Db 518 AspGlnLeuLeuGluLeuLeuArgAlaThrGlyAsnGluArgThrMetValPheValGlu 537  
 QY 1645 ACTAAGAAAGAGAGATTTTACGCAACTTTTCTTGTCCAGAAAGAAATATCACTACA 1704  
 Db 538 ThrLysArgSerAlaAspPheIleAlaThrPheLeuGlyGlnLysAlaLeuSerAspPheArgLeu 557  
 QY 1705 AGTATCCATGATGATCGGAACAGAGACCGGAGCAAGCTCTTGAGATTTTCGCTTT 1764  
 Db 558 SerIleHisGlyAspArgGluGlnArgGluArgGluLysAlaLeuSerAspPheArgLeu 577  
 QY 1765 GGAAGTCCGAGTCTTGTGCTACTCTCACTAGTACCTCCAGAGGCTGATATTGAAAT 1824  
 Db 578 GlyHisCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluGln 597  
 QY 1825 GTGCAATGTTATCAATTTTGAATTTTCTCTTCACTGATGATGATGATGATGATGAT 1884  
 Db 598 ValGlnHisValValAlaAsnPheAspMetProSerSerIleAspGlyTyrValHisArgIle 617  
 QY 1885 GGGCGTACTGCTGTTGTGGAAATATCTGCGACAGCAATTTCTTTTGAATCTTGAATCG 1944  
 Db 618 GlyArgThrGlyArgCysGlyAsnThrGlyArgAlaValaSerPhePheAsnProGluSer 637  
 QY 1945 GATPACCATTTAGACACGCTCTAGTAAAGTATGACAGATGCTCAACAGAGATTTCT 2004  
 Db 638 AspThrProLeuAlaArgSerLeuValLysValLeuSerGlyAlaGlnGlnValValPro 657  
 QY 2005 GCATGTGGAAGAAATGCTTGTAGTACATCACTCTGCTCAGAGGT---AGTACA 2061  
 Db 658 LysTyrLeuGluGluValAlaPheSerAlaHis-----GlyThrThrGlyPheAsnPro 675  
 QY 2062 AGAAGAAAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2121  
 Db 676 ArgGlyLysValPheAlaSerThrAspSerArgLysGlySer-----690  
 QY 2122 GGGTTTCTCTTCACAGAGCTCC-----AATCCA 2151  
 Db 691 ---PheLysSerAspGluProProProSerGlnThrSerAlaProSerAlaAlaAla 709  
 QY 2152 GTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2172  
 Db 710 AlaAspAspGluGluTyrGlu 716

RESULT 4  
 Q8QFUD PRELIMINARY; PRT: 715 AA.  
 AC Q8QFUD;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Vasa-like protein.  
 GN VEG.  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NC NCB1\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vatlani A., Olsen L.C.;  
 RT "Establishment of transgenic zebrafish lines expressing green

RT fluorescent protein in their germ cells.  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ111625; CAC84069.1; -  
 SO SEQUENCE 715 AA; 76880 MM; 3CE4E03F02073BA4 CRC64;

## Alignment Scores:

Pred. No.: 3.08e-147 Length: 715  
 Score: 1995.50 Matches: 426  
 Percent Similarity: 66.88% Conservative: 89  
 Best Local Similarity: 55.32% Mismatches: 150  
 Query Match: 50.47% Indels: 105  
 DB: 13 Gaps: 21

US-09-714-865-15 (1-2172) x Q8QF00 (1-715)

```

QY 10 GAAGATTGGAGAGCAGCAATCAACCTCATATGCTTCTTATGTTCCCAT- 60
DB ::::::::::::::::::::
QY 61 -----TTTGAGAGGATAGTATTCTTGAGAAAATGAGACAAAT 99
DB 22 Serannglyseraapgllyphe---lyserpetherthrglyalaglyasnaaplye 40
QY 100 TTTACAGGACTCCAGCTTCATCTCAAGAAATGATGACCTTCTCGAAGATCAT 159
DB 41 -----SeransergllygltthrgllyserSerTrpys----- 52
QY 160 TTCAATGAAAGTGATTTGCTGGCGGAGATTGGAAAACAGATGCTGTGACTGT 219
DB 53 -----Methnrglyaspserrpetharglyarg----- 61
QY 220 AATAAGCAGATATATACATCCCAATGGGTGTTTGAGATTGGAAGATTGGAAAC 279
DB 62 -----Glyglyarggly---glyserargglyglyarg 71
QY 280 AGAGCTTTTCA-----AACAGAGCTTTGAAAGATGATGCTTCTGTTCTGAGA 333
DB 72 GlyglyphesergllypelysSergllyleasplunsnllyseraapgllyglttrp--- 90
QY 334 GAGCTGTATGATGCTCGAAGATATCCAAACAGCAAGAGAGGTTTCCAGAGAGGC 393
DB 91 -----Asnnglyglyluserargllyargllyargllyglypne-----Arggly 105
QY 394 GGCATATGAGATGGAATTAATTCAGAA-----GCT 423
DB 106 Glyphesergllyserargspglunasnngllyasnaapargllytrpys 125
QY 424 TCAGGGCCATACAGAAAGAGTGAAGGTAAGTTCCGAGGTTGC---CGTGAAGATTT 480
DB 126 Glyglyluserargllyargllyargllypheglylyserpneargllyglyphe 145
QY 481 -----GCTAGAGTCCAAATTAATGACTTA 507
DB 146 Argapergllyasnngllyasnthrgllyargllypheglylunasnngllyasn 165
QY 508 GACCCAGCAGATGTATGACG-----CGCACTGGTGGCTTTTGGTTGTAGA 555
DB 166 Glyasnnapergllyglylunngllyargllyargllypneargllyglyphe 185
QY 556 AGACCACTATTAACTGACACAGTATGCTGATCTTCTCAAGCAGAAAGTGCAGTGA 615
DB 186 Arg-----Aspglyglyllyasnngllysergllylyasrg-----Gly 197
QY 616 AGTGAACGAGTGGTTACAAAGGTTTAAATGAGAAATTAATACAGGCTCTGGAAGAAT 675
DB 198 PhegllyargllyglyllypneargllyargasnngllylvalpheSer---LysvalThrThr 216
QY 676 TCTTGAAGTCAAGAGCAGAGAGGAGAAAGTATGATGATCTCAAGGACCAAAAGTGACC 725
DB 217 Alaasprylyleuasnngllyglyllysergllyasnala-----Glytrpvalval 233
QY 736 TACATACCCCTCTCTCACTGAGATGAGACTCCATCTTTGCACTTATCAAGACGCG 795
DB ::::::::::::::::::::

```

```

DB 234 TyrvalProProProProglugluserSerilepbeserhstryalathrgly 253
QY 796 ATAACTTCGACAAATACGACACTATTTCTTGGAAGTGTGACATGACACCA 855
DB 254 IleasnPhesaplystryaspaepileleuvalaspyalserllyserasnProProlys 273
QY 856 GCAATTCGACTTTTGAAGAGCTAACTCTGTGACACACTGAATACCAATTCGCTAA 915
DB 274 AlailemethrphedglunngllyllyleuCyaspserserleuserlyasnValserlye 293
QY 916 GCTGCTTATCTAGCTTACTCTCTGCAAAAATACAGTATTTCTCATCTACTGACGA 975
DB 294 SerGlyTyrVallyserProthrProvalGlnlyshiglylleProilleleSerlly 313
QY 976 CGAGATTGATGCTGTGTGCTCAAAAGGCTGGGAAGACTCGGCTTTTCTTACCA 1035
DB 314 ArgaspleumethalacyasalaglnnrglyserGlyyethrAlaAlaPheleuPro 333
QY 1036 ATTTGGCTCATATGATGATGATGAAATACCTGCACTGTTTAAAGATTGACGAA 1095
DB 334 IleleuGlnargPheethrthraepgllyvalAlaAlaSerlyserphesergllyumetGlnlu 353
QY 1096 CCAAGTGTATTTATGAGCACTCACTCGAAGATTTGCTCAACCAATTTATTTGGAAGCC 1155
DB 354 ProgluAlailellevalaProthrarggluleuileasnngllyletryleuGlnua1a 373
QY 1156 AGAAAATTTCTTTGGAGCTGTGTGATGAGAGCTGTGTATATATGAGGGGAACCCAGCTG 1215
DB 374 ArglysphehalatylglythrcysvalargProvalvalValtyrlyglylleasnthr 393
QY 1216 GCACTTCATTCATTCGACAAATAGTACAGGCTGTATATATATGCTACTCTCTGAGAA 1275
DB 394 GlyTyrThrleargglunvalleuLyssglyCysasnvalleuCyasbalthrProglyarg 413
QY 1276 CTGATGATATCTATAGGCAAAAGAAAGATTGGCTCAAAACATCAAAATCTTATGTTG 1335
DB 414 LeuileaspleuileglyllyargllyllyleGlyleuserlyvalargllyrleuvalleu 433
QY 1336 GATGAAGCTGATGCGATGTTGATGATGAGTTTGGTCCGAAATGAGAAAGTAAATTTCT 1395
DB 434 AspglunAlasprargmetleuasnmetglyllypneglunProglumetarglylsleuvala1a 453
QY 1396 TGCCAGGAATGCCATCAAGAGCAGCGCAAAACCTTATGTTCAAGCACTTTTCA 1455
DB 454 SerProgllymerProserlysglunarglunThrleuemerpheserlathrTyro 473
QY 1456 GAGGAATTCAAAGCTTGCTGACAGCTTTTAAATCAAAATTTATGTTGTTGCTGTT 1515
DB 474 GluasplileglunargmetAlaAlaasppheleuLyvalaspyrillepheleuAlaVal 493
QY 1516 GGCAAGTGGTGAGACATGTAGAGATGTTACAGACACCGTCTCCAGTTGGCCAGTTC 1575
DB 494 GlyvalvalGlyglyAlacyserasppvalGlnGlnThrilevalGlnvalaspglnlyr 513
QY 1576 TCAAAAAGAAAAAGCTCGTGAATTTCTGCAAAATAGGGAGATGAAAGAACTATGCTC 1635
DB 514 SerlyargAspGlnleuenglunleuargAlaThrGlyAsnngllyrghrmetVal 533
QY 1636 TTTGTTGAAATCAAGAAAAAGACAGATTTTACTGCACTTTTCTTTGTAAGAAAAAATA 1695
DB 534 PhevalGlnThrlyserSerAlasppheileAlaThrPheleuysglnngllyrlyle 553
QY 1696 TCAACTACAGATATCATGATGATGCGGAAACAGAGACGGGAGCAACTCTTGAGAT 1755
DB 554 SerThrThrSerillehlsolysasppargglunngllyargllyllyvalleuSerAsp 573
QY 1756 TTTGCTTTGAAAGTGCACAGTCTTTGTTGCTACTTCAAGTGCAGAGGCTGAT 1815
DB 574 PheargllyleuGlyhiesyProvalleuvalAlaThrserValAlaAlaargllyleuAsp 593
QY 1816 ATTGAAATGTCACACATGTATCAATTTTGAATCTTCTTCAACATTTGATGATATGTT 1875
DB 594 IleglunGlnvalGlnhlsvalvalasnPhesapmetProserSerilleaspgllyrval 613

```



QY 1876 CATGGAATGGGGCTACTGTCGTTGGTAATGACGAGCAATTCCTTTTGTAT 1935  
 |||||  
 Db 614 HieKrglIeglyagthrIyargCyGlyAsnThrIyargAlaValSerPheAsn 633  
 |||||  
 QY 1936 CTGGAATGGAATCAACCTTTGACGAGCCTCTAGTAAAGTATTGACAGATGCTCAACG 1995  
 |||||  
 Db 634 ProGluSerAspThrProLeuAlaIargSerLeuValIySerGlyAlaGln 653  
 |||||  
 QY 1996 GATGTCCTGATGCTGGAAAGAAATTCCTTGTACATACATCTCTGCTTACAGT 2055  
 |||||  
 Db 654 ValValProIyStrIleuGluValAlaPheSerAlaHis-----GlyThrThrGly 671  
 |||||  
 QY 2056 ---AGTACAGAGAAAGCTGTTTGCATGATGATACAGAAAGGCAAGACATTTG 2112  
 |||||  
 Db 672 PheAsnProIyGlyValPheAlaSerThrAspSerIyGlyGlySer----- 689  
 |||||  
 QY 2113 AACACAGCTGGTTTCTTCTTACGAGCTCC----- 2145  
 |||||  
 Db 690 -----PelySerAspGluProProSerGlnThrSerAlaProSerAla 705  
 |||||  
 QY 2146 ---AATCAGTAGATGATGATGATGAT 2172  
 |||||  
 Db 706 AlaAlaAlaAlaAspAspGluIyPglu 715  
 |||||  
 RESULT 5  
 091372 PRELIMINARY; PRT; 700 AA.  
 AC 091372;  
 DT 01-NOV-1996 (TREMURel. 01, Created)  
 DT 01-JUN-1998 (TREMURel. 06, Last sequence update)  
 DT 01-MAR-2002 (TREMURel. 20, Last annotation update)  
 DE DEAD box protein.  
 GN VGL.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94200507; PubMed=8150200;  
 RA Komiyama T., Itoh K., Ikenishi K., Furusawa M.;  
 RT "Isolation and characterization of a novel gene of the DEAD box  
 RT protein family which is specifically expressed in germ cells of  
 RT Xenopus laevis".  
 RL Dev. Biol. 162:354-363(1994).  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RA Komiyama T.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
 DR EMBL; AF046043; AAC03114.1; -.  
 DR HSP; Q58083; IHV8.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD box.  
 DR InterPro; IPR001650; Helicase C.  
 DR InterPro; IPR000685; Rubisco\_Large.  
 DR Pfam; PF00270; DEAD. 1.  
 DR Pfam; PF00271; Helicase\_C. 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC\_C; 1.  
 DR PROSITE; PS00039; DEAD ATP HELICASE; 1.  
 DR PROSITE; PS00157; RUBISCO\_LARGE; UNKNOWN\_1.  
 KW ATP-binding; Helicase; RNA-binding.  
 SQ SEQUENCE 700 AA; 78240 MW; C504ECA38EFB0B7E CRC64;

## Alignment Scores:

Pred. No.: 1.98e-144 Length: 700  
 Score: 1959.50 Matches: 409  
 Percent Similarity: 68.17% Conservative: 105  
 Best Local Similarity: 54.24% Mismatches: 153  
 Query Match: 49.56% Indels: 87

DB: 13 Gaps: 18  
 US-09-714-865-15 (1-2172) x 091372 (1-700)  
 QY 7 GATGAATTTGGAGACGAAATCAACCCATATGCTTCTATGTTCCCATATTGAG 66  
 |||||  
 Db 2 GluGluAsnTrpAspThrGluIleGluThrGluLysProThrIyValProAsnPh 20  
 |||||  
 QY 67 AAGGATAGTATTTCTGGAAATGAGCAATTTTAAAGAGACTCAGCTCATCATCA 126  
 |||||  
 Db 21 ---SerThrIeuGluThrGluAsnThrAspAsnTr-----SerAlaTrSerAsn 37  
 |||||  
 QY 127 GAATGATGATGAGACCTTCTCGAAGAGATCATTCATGAAGATTTGCTCGG 186  
 |||||  
 Db 38 AspIleAsnAsn-----GlnAsnTrAspSerGlu 47  
 |||||  
 QY 187 CGGAATTTTGGAAACAGAGATGCTGAGTGAATTAAGGAGATTAATCATCCACATG 246  
 |||||  
 Db 48 ArgSerPheGlyAsnArg---GlyGlyTyraIySerGluArgSerProSerAsnPh 66  
 |||||  
 QY 247 GGTGTTTGGAGTTGGAAGAGTTTGGAAACAGAGTTTTCAAACAGAGTTGAA 306  
 |||||  
 Db 67 -----AsnArgIy-----SerArgThrGlu 73  
 |||||  
 QY 307 GATGATGATGATCTGCTGTTCTCGAGAGAGTCTAGTATGATGCTC----- 351  
 |||||  
 Db 74 ArgGlyArgGlyArgGlyPheGlyThrAsnArgAsnAspAsnTrSerSerGluArgAsp 93  
 |||||  
 QY 352 -----GAAGTAAATCCACACGAGACAGGCTTTCCAAAGAGCGGCTATCGA 402  
 |||||  
 Db 94 ValPheGlyAspAspGluArgAspGluArgArgGlyPheProGlyArgGlyTyra 113  
 |||||  
 QY 403 GATGAATTAATTCAGAGCTTCAGAGCCATACAGAAAGAGTGAAGTATTCGA 462  
 |||||  
 Db 114 GlyAsnGluAspGlyGlnLysProAsnAlaPheArg-----GlyArgGlyPheArg 131  
 |||||  
 QY 463 -----GTTGCCGTGAGAGATTTGGTCTAGAAAGTCCA 495  
 |||||  
 Db 132 AsnGluAsnGluGluAsnArgGlyPheGlyGluArgGlyGlyPheArgSerGluAsnGly 151  
 |||||  
 QY 496 AATTAATGACTTAAGACCCAGACGATGATGACGAGCTGTCCTTTTGGTCTAGA 555  
 |||||  
 Db 152 GluArgAsnPhAsp-----AsnArgIyAspPheGlyAsnSer 164  
 |||||  
 QY 556 -----AGACCAGTATTAAGTGGACAGCT-----AATGGATGACT 591  
 |||||  
 Db 165 GlyGluGluGluAsnArgProArgSerTyraIyArgGlyGlyPheAsnAsnSerAspThr 184  
 |||||  
 QY 592 TCT-----CAAGCAGAGAGTGCAGTGAAGTGAAGAGGTGTTACAAAGT 639  
 |||||  
 Db 185 GlyGlyArgGlyArgArgGlyGlyArgGlyGlySerGlnTrArgGlyTyraGly 204  
 |||||  
 QY 640 TTAATGAGAGTAAATTAACAGCTCTGGAAGAAATTTTGAAGTCAAGAGAGGA 699  
 |||||  
 Db 205 ArgAsnGluGluVal-----GlyValGlySerGlyLysSerGlnGluGly 220  
 |||||  
 QY 700 CGAAGAGTGTGATCTCAAGACCAAGTACCTACCTACCTCCACCTGAG 759  
 |||||  
 Db 221 AsnGlu---LysAspGluLysProLysLysValThrTyraIleProProProProAsp 239  
 |||||  
 QY 760 GATGAGACTCTCATCTTTGACATTTACAGAGGCAATTAAGTTCGAAATAGACACT 819  
 |||||  
 Db 240 GlyGluAsnTrpAsnThrPheArgGlnTyraGlnSerGlyIleAsnPhAspLysTyra 259  
 |||||  
 QY 820 ATTTCTTGAAGTGTCTGACATGATGACACCAAGCAATTTGACTTTGAAGAGCT 879  
 |||||  
 Db 260 IleLeuValAspValThrGlyLysAspValProProAlaIleLeuThrPheGluAla 279  
 |||||  
 QY 880 AATCTGCTGAGACGAGATTAACAGATGATGATGATGATGATGATGATGATGAT 939  
 |||||  
 Db 280 AsnLeuGluThrLeuArgArgValAlaArgAlaGlyTyraIyValLysLeuThrPro 299  
 |||||  
 QY 940 GTGCAAAAATACAGTATTCCTATCATCTTGCAGAGAGAGATTTGATGCTGTCAA 999  
 |||||

```

Db 300 ValGlnYshSserIleProIleIleMeclalaglyrghasrleuMetclalacYalagln 319
Qy 1000 ACAGGGCTCGGAGAGACTCGCGCTTTCTCTACCAATTTGGCTCATATGATGCATGAT 1059
Db 320 ThrlYserGlyYshThrlAlaAlaPheleuProIleleuSerTYrMetMeclasp 339
Qy 1060 GGAATACCTGCACGCTTTTAAAGAGTTGCAGAGACCAAGCTGATATATATGTGACCA 1119
Db 340 GlyIleThrlAlaSerGlnTYrleuGlnleuGlnleuProGlnAlaIleIleIleIlePro 359
Qy 1120 ACTGAGAAATTTGTCACACAGATTTATTTGGACCCAGAAATTTCTTTGGACCTGT 1179
Db 360 ThrlArgIleuIleleuGlnIleTYrleuAspIleArgYshSerTYrGlyThrCys 379
Qy 1180 GTAGAGCTGTGTTATATATAGGGGAGACCAAGCTGGACATTCATATCCCAATATGTA 1239
Db 380 ValArgProValIleValTYrGlyGlyIleGlnProValIleIleIleIleIleIleIle 399
Qy 1240 CAAGGCTGTATATATATATGCTACTGCTGAGAGACTGATGATCATATGCAAGAA 1299
Db 400 LysGlyCysAsnIleleuCysAlaThrProGlyArgleuAspIleValSerLYsGln 419
Qy 1300 AAGATTGCTCAACACAGATCAAACTACTATTGTTGATGAGCTGATCGCATGTGGAT 1359
Db 420 LysIleGlyleuSerLYsleuArgTYrleuValleuAspGlnIleAspArgMetleuAsp 439
Qy 1360 ATGGCTTTGCTCCAGAAATGAGAAGTTATTTCTTCCAGAAATGCCATCAAGGAA 1419
Db 440 MetIlyPheAlaProGlnIleGlyleuMetThrlYshProGlyMetProThrLYsGln 459
Qy 1420 CAGGCGCAAAACCTTATGTTGATGTCACACTTTCCAGAGAAATTCAGAGTTGGCTGCA 1479
Db 460 LysArgGlnThrleuMetPheSerAlaThrTYrProGlnGlnIleArgArgleuAlaSer 479
Qy 1480 GAGTTTTTAAATCAAAATATCTGTTGTTGCTGTTGTCAGACAGTGGTGACATGTAGA 1539
Db 480 AsnTYrleuYshSerGlnIleleuPheValIleValGlyleuValIleGlyAlaCysSer 499
Qy 1540 GATGTTGACGAGACCGCTTCCAAAGTTGGCCAGTTCTCAAAAAGAGAAACCTGTTGAA 1599
Db 500 AspValIleGlnThrValleuGlnMetArgGlnleuGlnIleYshMetGlnYshleuGln 519
Qy 1600 ATTCTGGAAACATATAGGGGATGAGAAAGCTATGCTTTGTTGAAATCAAAAGAAAGCA 1659
Db 520 IleleuYshSerSerGlnYshGlnArgThrleuIlePheValAsnThrLYsLYsValAla 539
Qy 1660 GATTTTACTGCACATTTTCTTTGTCAGAAAGAAATATCAAGTACAGATGCATGAT 1719
Db 540 AspheIleIleIleIleGlyTYrleuCysGlnIleYshPheSerSerThrIleIleIleIle 559
Qy 1720 CGGAAACAGAGAGCGGAGCAAGCTTTGGAATTTTCCCTTTGGAATGCCCAAGT 1779
Db 560 ArgGlnGlnTYrGlnArgGlnIleuSerIleuThrAspPheArgThrGlyLYsCysThrVal 579
Qy 1780 CTTGTGCTACTTATAGTATGTCAGAGGCTGATATGAAATGTCAGACATGTTATC 1839
Db 580 IleValCysThrAlaValAlaIleArgGlyleuAspIleGlnIleValGlnIleValIle 599
Qy 1840 AATTTGATCTTCTTACCATGATGATGATATGTTGATTCGAATTTGGCGCTATGCTGT 1899
Db 600 AsnTYrAspValProLYsGlnValAspIleuTYrValIleIleIleIleIleIleIleIle 619
Qy 1900 TGTGGAAATCTGCGACAGCAATTTCTTTTGTGATCTTGAATGGATACCATTTAGCA 1959
Db 620 CysGlyAsnThrGlyLYsAlaThrSerPheAsnValGlnIleAspAsnIleValIleAla 639
Qy 1960 CAGCTCTAGTAAGATTTGACAGATGTCACAGATGTTCTGCAATGTTGGATGGAAGAA 2019
Db 640 ArgProleuValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 659
Qy 2020 ATTGCTTTTACTATACATCTCTGCTTCAAGTGTGATGACAGAAACGTGTTTGA 2079

```

```

Db 660 IleAla-----PheGlyGlyIleGlyAlaIleuAsnSerPheTYr 672
Qy 2080 TCAGTTGATACC-----AGAAAGCGAAGAGCACTTTGAACACAGCTGGTTTCT 2130
Db 673 AlaAlaAspSerMetGlyGlnIleGlyAlaIleValAlaIleThrProSerPheAla 692
Qy 2131 TCTTCAAGAGCTCCCATCCATGATGATGATGATGATGATGATGATGATGATGAT 2172
Db 693 GlnGln-----GlnGlnAlaSerThrAsp 700

RESULT 6
Q902F6 PRELIMINARY; PRT; 617 AA.
AC Q902F6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE VASA.
GN OLVAS.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphna; Acanthopterygii; Percormorphna; Atherinomorphna;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinomura A., Tanaka M., Kobayashi T., Nagahama Y., Hamaguchi S.;
RT "The vasa-like gene, olvas, identifies the migration path of
RT primordial germ cells during embryonic body formation stage in the
RT medaka, Oryzias latipes."
RL Development 42:317-326(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21126920; PubMed=11226275;
RA Tanaka M., Kinoshita M., Kobayashi D., Nagahama Y.;
RT "Establishment of medaka (Oryzias latipes) transgenic lines with the
RT expression of green fluorescent protein fluorescence exclusively in
RT germ cells: A useful model to monitor germ cells in a live
RT vertebrate."
RL Proc. Natl. Acad. Sci. U.S.A. 98:2544-2549(2001).
DR EMBL; AB063484; BAB61047.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; UNKNOWN 1.
KW ATP-binding; Helicase.
SQ SEQUENCE 617 AA; 66737 MW; C1B8A3462760BA50 CRC64;

Alignment Scores:
Pred. No.: 9.27e-140 Length: 617
Score: 1899.50 Matches: 386
Percent Similarity: 74.45% Conservative: 86
Best Local Similarity: 60.88% Mismatches: 129
Query Match: 48.04% Indels: 33
DB: 13 Gaps: 13

US-09-714-865-15 (1-2172) x Q902F6 (1-617)
Qy 304 GAAGATGATGATGCTGCTGTTCTGAGAGAGCTAGTAAGCTGCAAGATTAATCCA 363
Db 6 GlnGlnIleuThrAlaProSerPheAlaProValSerSerThrAspAlaIle-----Pro 23
Qy 364 ACAAGCAACAGAGGCTTTTCCAGAGAGCGGCTATGACAT--GGAAATTAATTCAGAA 420
Db 24 GlnArgSer-----SerTrpAsnGlyGlyAlaArgAspSerGlyAsnAspGlyAsp 40
Qy 421 GCTTACAGGCGCATACAGAGAGGTGCAAGAGTATTC---CGAGTTCCCGTGAAGCA 477
Db 41 SerTrpAsnArgSerAsnArgGlyArgGlySerAlaIleArgIleGlyArgGlyGly 60

```

QY 478 TTGGTCTAGGAAGTCCAAATTAATGACTTAGACCCAGACGAATGATACGCGCATGCT 537  
 Db 61 ArgGlyArgGlyPheGlyArgSerAspGlnasp-----GluLeuAsnGly 75  
 QY 538 GGCCCTTTTGGTCTTAGAAGACAGTATTAAAGTGCACAGTAATGGTACTCTTCAA 597  
 Db 76 Gly---GlyGlyAspSerGluAsnGlyPheArgGlyArgGlyGlyArgGlyGly 94  
 QY 598 AGAGAAGTGCAGTGAAGTGAACGA-----GGTGGTTCACAAGGTTTAAATGAA 648  
 Db 95 PheArgSerGly---GlyGlyGlyArgGlyArgGlyGlyArgGlyGlyArgGlyGly 113  
 QY 649 GAAGTAATA---ACAGGCTCTGGAAGAAATTTGGAAGTGCAGAGCAGAGAGAGAA 705  
 Db 114 AspValPheAlaAlaGlyAspGlyArgGly-----AlaGluAsnSerAsp 128  
 QY 706 AGTAGATATCTACAGACCAAAAGTACCTTACCTCCCTCCTCCACCTGAGAGATGAG 765  
 Db 129 AlaAlaAspProGluArgProGlyValThrTyrIleProPheSerLeuProGluAspGlu 148  
 QY 766 GACTCCATCTTGGACATTAATGACAGACGCAATTAAGTTCAGCAATATGACACTATCTT 825  
 Db 149 AspSerIlePheSerHisTyrLysMetGlyIleAsnPheAspLysTyrAspAspIleLeu 168  
 QY 826 GTGGAAGTGTGACATGATGACACCAAGCAATTCGACTTTGGAAGAGCTATCTC 885  
 Db 169 ValAspValSerGlyThrAsnLeuProAlaAlaIleMetThrPheGluGluAlaLysLeu 188  
 QY 886 TGTACAGACTGTAATACCAACATTGCTAAAGCTGTTATTAAGCTTACTCTGTCGA 945  
 Db 189 CysGluSerLeuGluAsnIleSerArgSerGlyTyrValLysProThrProValGln 208  
 QY 946 AAATACAGTATCTTATCATCTTATGACAGACGAGATTGATGGTGGTCCAAACAGG 1005  
 Db 209 LysTyrGlyLeuProIleIleSerAlaGlyArgAspLeuMetAlaCysAlaGlnThrGly 228  
 QY 1006 TCTGGAAGACTGGGCTTTTCTCTCAATTTTGCTCATATGATGATGATGATGATA 1065  
 Db 229 SerGlyLysThrAlaAlaPheLeuLeuProIleLeuGlnGlnLeuMetAlaAspGlyVal 248  
 QY 1066 ACTGCCAGTGTGTTTAAAGCTGACAGAAACAGAGTATTAATTTAGACCAACTGCA 1125  
 Db 249 AlaAlaSerArgPheSerGluIleGlnGluProGluAlaValIleAlaIleProThrArg 268  
 QY 1126 GAATGGTCAACAGATTTATTTGGAAGCCAGAAATTTCTTTGGGACTGGTGTAGA 1185  
 Db 269 GluLeuIleAsnGlnIleTyrGlnGluAlaArgLysPheSerPheGlyThrCysValArg 288  
 QY 1186 GCTGTGTTTATATAGGGGGAACCCAGCTGGACATTCAATTGCAAAATAGTACAAGC 1245  
 Db 289 ProValValValTyrGlyGlyValAsnThrGlyTyrGlnMetArgGluIleGluLysGly 308  
 QY 1246 TGTAAATATATATGCTACTCTCTGGAAGACTGATGATATCATAGCCAAAGAAAGATT 1305  
 Db 309 CysAsnValLeuCysGlyThrProGlyArgLeuLeuAspMetIleGlyArgGlyLysVal 328  
 QY 1306 GGTCTCAACAGATCAATCTTATGTTGGATGAAGCTGATGCGCATTTGGATATGGCT 1365  
 Db 329 GlyLeuSerLysValArgHisLeuValLeuAspIleAlaAspArgMetLeuAspMetGly 348  
 QY 1366 TTGTGTCAGAAATGAAGATTATTTCTTCCAGAAATGACCATAAAGAACACAGC 1425  
 Db 349 PheGluProAspMetArgArgLeuValGlySerProGlyMetProSerLysGluGluArg 368  
 QY 1426 CAAACCTTATGTTCACTGACACTTTTCCAGAGAAATTCAAAGTTGGCTGACAGATT 1485  
 Db 369 GlnThrLeuMetPheSerAlaThrPheProGluLysPheGlnArgLeuAlaAlaAspPhe 388  
 QY 1486 TTAAGCAATATATGCTTTGTTGCTGTTGGAACAAGTGGTGGATGATGAGATGTT 1545  
 Db 389 LeuLysValAspTyrLeuPheValAlaValGlyValGlyGlyAlaCysThrAspVal 408  
 QY 1546 CAGCAGACCGTTCTCAAGTTGGCCAGTTCTCAAAAGAGAAAGCTGTTGAATCTG 1605

Db 409 GluGlnThrPheLeuGlnValThrLysPheAsnLysArgGluGlnLeuLeuAspLeuLeu 428  
 QY 1606 CGAAACCTAGGAGATGAAAGAACTATGCTTTGTTGAACCTAAGAAAGAGATTTT 1665  
 Db 429 ArgThrIleGlySerGluArgThrMetValPheValGluThrLysAspGlnAlaAspPhe 448  
 QY 1666 ACTGCAACTTTCTTTGTCAGAAAGAAATATCAACTACATGATTCATGGTGAATCGGAA 1725  
 Db 449 IleAlaAlaPheLeuCysGlnGluLysValProThrThrSerIleHisGlyAspArgGlu 468  
 QY 1726 CAGAGAGCGGAGACAGCTCTTGAAGATTTTCGCTTGGAAAGTCCCAAGTTCTTGT 1785  
 Db 469 GlnArgGluArgGluLysAlaLeuAlaAspPheArgSerGlyLysCysProValLeuVal 488  
 QY 1786 GCTACTTCACTACTGTCGACAGAGGCTCGGATTAATGAAATGCAACATGTTATATTT 1845  
 Db 489 AlaThrSerValAlaSerArgGlyLeuAspIleProAspValGlnHisValAlaAspPhe 508  
 QY 1846 GATCTTCTCTTCTACATGATGATGATATGTTCAATGCAATTTGGGCTGCTTGGGG 1905  
 Db 509 AspLeuProAsnThrIleAspAspTyrValHisArgIleGlyArgThrGlyArgGly 528  
 QY 1906 AATACCTGCGACGCAATTTCTTTTGTGATCTGGAATCGATACCATTTAGACAGCT 1965  
 Db 529 AsnThrGlyArgAlaValSerPheTyrAspProAspValAspSerGlnLeuAlaArgSer 548  
 QY 1966 CTAGTAAATATTTGACAGATGCTCAACAGAGATTTCTCCATGCTTGGAAAGAAATGCC 2025  
 Db 549 LeuValGlyLysLeuAlaLysAlaGlnGlnGluValProSerThrLeuGluGluSerAla 568  
 QY 2026 TTATGATACATACATTCCTGCTCAGTGTAGTACA-----AGAGAAACGTG 2073  
 Db 569 Phe-----GlyAlaHisGlySerAlaAlaPheAsnProSerGlyArgThr 583  
 QY 2074 TTTCATCAGTTGATACAGAAAGGCGACAGACCTTTGAACACAGCTGGGTTTCTTCT 2133  
 Db 584 PheAlaSerThrAspSerArgLysGlyLysSerPheGlnAspSerSerValLysThrGln 603  
 QY 2134 TCACAGACT---CCCAATCCAGTATGATGATGATGATGATGATGATGATGAT 2172  
 Db 604 ProAlaAlaProProAlaAlaAlaAspGluAspAspThrGlu 617  
 RESULT 7  
 Q9PT10 PRELIMINARY; PRT; 647 AA.  
 AC Q9PT10; 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Vasa.  
 GN VAS.  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVARY;  
 RX MEDLINE=20160526; PubMed=10694742;  
 RA Yoshizaki G., Sakatani S., Tomioka H., Takeuchi T.;  
 RT "Cloning and characterization of a vasa-like gene in rainbow trout and  
 its expression in the germ cell lineage."  
 RL Mol. Reprod. Dev. 55:364-371(2000).  
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
 DR EMBL; AB032566; BAA88059.1; -;  
 DR HSSP; Q58083; IHW8.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000628; DEAD\_box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD\_1.  
 DR Pfam; PF00271; helicase\_C\_1.

DR SMART; SMO0487; DEXDC; 1.  
 DR SMART; SMO0490; HELIC; 1.  
 DR PROSITE; PS00039; DEAD ATP HELICASE; 1.  
 KM ATP-binding; Helicase; RNA-binding.  
 SQ SEQUENCE 647 AA; 68525 MW; 88F25AE7A68F51C2 CRC64;

## Alignment Scores:

Pred. No.:	1,27e-138	Length:	647
Score:	1885.00	Matches:	395
Percent Similarity:	68.32%	Conservative:	86
Best Local Similarity:	56.11%	Mismatches:	141
Query Match:	47.67%	Indels:	82
		Gaps:	15

US-09-714-865-15 (1-2172) x Q9PT10 (1-647)

```

QY 79 TCCTGAGAAATGAGACAAATTTAAACGACCTCCAGCTTCATCATGAAATGATGAT 138
   |||||
Db 20 SerGlyGlnSerSerPheGlyArgProThrAspLysValSerSerTyrAsnSer 39
   |||||
QY 139 GGACCTTCTCGAAGAGATCATTTGATGAAAGTGGATTTGCTGGCGGAAATTTTGA 198
   |||||
Db 40 Gly-----CysGlyGlyPheGlyGlyGlyArgGly 51
   |||||
QY 199 AACAGAGATGCTGTGATGATTAATTAAGCAGATTAATCATCCAAATGGGTGTTTGA 258
   |||||
Db 52 SerArg-----GlyGlyGly 57
   |||||
QY 259 GTTGGAAAGACTTTTGGAAACAGAGCTTTTCAACACAGAGTTTGAAGATGGATGAC 318
   |||||
Db 58 AspPheLysSerPheSerSerGlyAlaGlyLysAsnGlyAsn---GluAspLysAspSer 76
   |||||
QY 319 TCTGTTTCTGAGAGAGTCTAGTAATGATGCTGCAAGATATCCAAACGAGACAGAGG 378
   |||||
Db 77 Ser-----Tyr----- 78
   |||||
QY 379 TTTTCCAGAGAGCGCGCTATCGAGATGAAATTAATTCAGAGCTTCAGGCGCATACAG 438
   |||||
Db 79 -----AenSerGlyGlyPheArg 85
   |||||
QY 439 AGAGGTGAAGAGGTAGTTCCGAGCTTCCGCTGAGAGATTTGGTCTAGAGATCCAAAT 498
   |||||
Db 86 GlyArgGlyArgGlyGlyArgGlySerArgGlyGly-----GlyLLeArgAsn 102
   |||||
QY 499 AATGATTTAGACCCAGACGATGATGACGACGACCTGGTCTTTTGGTCTAGAGAA 558
   |||||
Db 103 GlyAspAspGlyTyrAsp-----GlyGlyPheGlyGlySerGln--- 115
   |||||
QY 559 CCAGTAATTAAGTGCACAGGT-----AATGTGATACTTCTCAAGCAGAAAGTGGAGT 612
   |||||
Db 116 -----GlyGlyArgGlyGlyArgGlyPheArgSerGlyGlyAspGlyGly 132
   |||||
QY 613 CGAGGTGAAGAGGTGTTTCAAAAGTTTAATGAAGAAATTAACA---GGCTTGGGA 669
   |||||
Db 133 GlyPheGlyGlyGlyGlyTyrArgGlyArgAspGlnGlnLLePheSerLysGlySerThr 152
   |||||
QY 670 AAGATTTCTTGAAGTGAAGAGCAGAAAGAGAGAAAGTACTATCTCAAGACCAAAA 729
   |||||
Db 153 MetAspGly-----GlyGlyAspGlyGlyAsnProGlyProProLys 166
   |||||
QY 730 GTGACCTACATACCCCTCTCTCCACCTGAGAGATGAGATCCATCTTGGACATTTACG 789
   |||||
Db 167 ValThrTyrValProGlnAlaLeuGlyLysGlnSerSerLLePheAlaHisTyrGln 186
   |||||
QY 790 ACAGGCAATTAATCTCGAATTAAGACATTTCTTGGAAAGTCTGAGACATGATGCA 849
   |||||
Db 187 SerIlyIleAsnPheAspLysTyrAspAspLLeLeuValAspValSerGlySerAspPro 206
   |||||
QY 850 CCAGCAGCAATTTGACCTTTGAGAGAGCTATCTCTGTCAGACATGAAATTAACAAT 909
   |||||
Db 207 ProLysAlaIleMetGlyPheGlnGlnAlaAlaLeuGlySerGlnSerLLeAsnArgAsnIle 226
   |||||
QY 910 GCTAAAGCTGGTTTACTTAAGCTTACTCTGTGCAAAAATTAACAGTATTCCTATCACTT 969

```

```

Db 227 SerLysSerGlyTyrLysLysProThrProValGlnLysHisGlyLLeProLLeIleAla 246
   |||||
QY 970 GCGAGCAGACATTTGATGGCTTGGCTGCACAAACGGGCTGGGAAAGCTGGGCTTTTCTC 1029
   |||||
Db 247 AlaGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlySerIlySerThrAlaAlaPheLeu 266
   |||||
QY 1030 CTACCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1089
   |||||
Db 267 LeuProLLeLeuGlnGlnLeuMetValAspGlyValAlaAlaSerGlnPheSerGlnLLe 286
   |||||
QY 1090 CAGAACCCAGAGTGTATTTATTTAGCACCAACTCGAAGATTTGGTCAACGATTTATTTG 1149
   |||||
Db 287 GlnGlnProGlnValIleIleValAlaProThrArgGlnLeuLLeAsnGlnLLeTyrMet 306
   |||||
QY 1150 GAAGCCGAGAAATTTTCTTTGGGAGCTGTGTAAGCTGTGTGTTATATGAGGGGAACC 1209
   |||||
Db 307 GlnAlaArgLysPheAlaHisGlyLysThrCysValArgProValAlaValIlyGlyGlyLLe 326
   |||||
QY 1210 CAGCTGGACATTCATTCGACAAATAGTCAAGAGCTGTAAATATATATGCTACTCTCT 1269
   |||||
Db 327 SerThrGlyHisThrIleArgGlnLLeLeuLysGlyCysAsnValLeuGlySalThrPro 346
   |||||
QY 1270 GGAAGCTGATGATATCATAGCCAAAGAAAGAAATTTGCTTCAACAGATCAAAATCTTA 1329
   |||||
Db 347 GlyArgLeuMetAspLLeIleGlyArgGlyLysGlyLLeGlyLeuSerLysLeuArgTyrLeu 366
   |||||
QY 1330 GTTTTGGATGAAGCTGATGCGCATGTTGGATATGAGGTTTGGCTCAGAAATGAAGAGTTA 1389
   |||||
Db 367 ValLeuAspGlnAlaAspArgMetLeuAspMetGlyPheGlnProAlaMetAlaArgLysLeu 386
   |||||
QY 1390 ATTTCTTGGCCAGAAATGCCATCAAGGAAAGACAGCGCCAAACCTTATGTTCACTGCAACT 1449
   |||||
Db 387 ValGlySerProGlyLLeProAlaLysGlnAspArgGlnThrLeuMetPheSerAlaThr 406
   |||||
QY 1450 TTTCCAGAGAAATTCAAAGCTTGGCTGCAGAGTTTAAAGTCAAAATTAATCTGTTGTT 1509
   |||||
Db 407 TyrProGlnAspLLeGlnLysLeuAlaAlaAspPheLeuLysAspTyrLeuPheLeu 426
   |||||
QY 1510 GCTGTTGGAACAAGTGGGTGAGCATGATGATGATGATGATGATGATGATGATGATGATG 1569
   |||||
Db 427 AlaValGlyValValGlyGlyAlaCysSerAspValGlnLysValValGlnValThr 446
   |||||
QY 1570 CAGTTCTCAAAAAGAGAAAGCTGTTGAATTTCTCGAAACATAGGAGGATGAAGAACT 1629
   |||||
Db 447 LysPheSerLysArgAspGlnLeuLeuGlnValLeuLysThrArgLysSerGlnArgThr 466
   |||||
QY 1630 ATGCTCTTTGTAACCTAAAGAAAGAAAGCAGATTTTACTGCAACTTTCTTTGTCAGAA 1689
   |||||
Db 467 MetValPheValGlnThrLysArgGlnAlaAspPheIleAlaThrPheLeuGlyArgGln 486
   |||||
QY 1690 AAAATATCAACTCAAGTATCCATGATGATGATGATGATGATGATGATGATGATGATGATG 1749
   |||||
Db 487 LysValAsnThrThrSerLLeHisGlyAspArgGlnGlnArgGlnGlnGlnAlaLeu 506
   |||||
QY 1750 GGAGATTTTGGCTTGGAAAGTCCAGATTTCTTGTCTACTTCACTAGTGCAGAGGG 1809
   |||||
Db 507 GlyAspPheArgSerGlyArgCysProValLeuValAlaHisSerValAlaAlaArgGly 526
   |||||
QY 1810 CTGATATTTGAAATGTGCAACATGTTATCAATTTTGTATTTCTTCTTCACTGATGAA 1869
   |||||
Db 527 LeuAspLLeLysAspValGlnHisIleValAsnPheAspLeuProAsnLLeAspGln 546
   |||||
QY 1870 TATGTCATGATGATGGGCTGATGCTGCTGTTGGGAAATATCTGGCAGAGCAATTCCTTT 1929
   |||||
Db 547 TyrValHisArgLLeGlyArgGlnArgCysGlyAsnThrIlyArgAlaValCysPhe 566
   |||||
QY 1930 TTTGATCTTGAATCGATATAACATTTAGCACACAGCTCTAGTAAAGTATTTAGACATGCT 1989
   |||||
Db 567 PheAspProGlyAlaAspGlyAsnLeuAlaArgSerLLeuValLysValLLeuSerGlyAla 586
   |||||
QY 1990 CAAAGAGATTTCTCATAGTGTGGAAGAAATTTGCTTATGATCATCAATTCCTGCTTC 2049
   |||||

```



```

Db      406 Alacysenaspvalglnthrphevalglnvalthrlysphealalyarglgn 425
Qy      1591 CTGTCGAATTTCTGCCAAACATAGGGAGTAAAGAACTATGGCTTTGTTGAACTAAG 1650
Db      426 Leuileaspneuylserthrnglyserglarthrvalphevalglutrllys 445
Qy      1651 AAAAAGCAGATTTTACTGCACCTTTCTTGTCAACAAAAATATCACTACAGATATC 1710
Db      446 Arglnalasphelealathrleleucyglndlysrpercthrserille 465
Qy      1711 CATGGTATCGGAGAACAGAGACCGGAGCAAGCTTTGAGATTTTCCTTTGGAAG 1770
Db      466 Hieglyasparglgnthrpglnarglgnlhalaleuglyspheargserglylys 485
Qy      1771 TGGCCAGTTCTTGTCTACTTCACTAGTACCTCCAGAGGCTGATATGAAATGTCGA 1830
Db      486 Cyservalneuvalalathrservalglvalarglyleuaspillepropaspvalgln 505
Qy      1831 CATGTTCAATTTGATCTCTCCCTTACCATGATGATATGTTTCATGAAATGGGCGT 1890
Db      506 Phevalvalenpheasphelepropasnamlleaspilutryvalhlsarglleglyarg 525
Qy      1891 ACTGTCGTGTGGGAATCTGCGAGAGCAATTTCTTTTGTATCTTGAATCGATTAAC 1950
Db      526 Thrglyargcysgllyasnthrnglyargalvalserpetherlyaspseroglyalaspser 545
Qy      1951 CATTTACACAGCCTCTAGTAAAGTATTTGACAGATCTCAACAGATTTCTGCTGATGG 2010
Db      546 Glnduulayrserleuvalthrleleuserlylainglnglualprosertrp 565
Qy      2011 TTGGAAGAAATTTGCTTATGACATACATTCCT---GGCTTACGTGTAGTACAGAGGA 2067
Db      566 Leuglnduulhealaspsergllyhlsrproserglpheasnpropolarglyhls 585
Qy      2068 AACGTGTTTCATCATGTTGATACCAAGAGGCAAGCACTTTGAACACAGCTGGGTTT 2127
Db      586 -----Phealaserthraspheularglyproglin-----GlyglyPhe 599
Qy      2128 -----TCTTCTTCACAGCTCCCAATCCAGTACAGATGATGATCA 2166
Db      600 Pheglaspgllyservalthrserglnduulalaglthrvalalaspaspglgn 619
Qy      2167 TGGGAT 2172
Db      620 Trpgln 621

```

## RESULT 9

```

Q9DEG3  Q9DEG3  PRELIMINARY;  PRT;  645 AA.
AC  Q9DEG3;
DT  01-MAR-2001 (Tremblrel. 16, Created)
DT  01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT  01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE  Vasa.
GN  VAS.
OS  Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC  Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC  Cichlidae; Oreochromis.
OX  NCBI_TaxID=8128;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20544961; PubMed=11091081;
RA  Kobayashi T.; Kajitara-Kobayashi H.; Nagahama Y.;
RT  "Differential expression of vasa homologue gene in the germ cells
RT  during oogenesis and spermatogenesis in a teleost fish, tilapia,
RT  Oreochromis niloticus."
RL  Mech. Dev. 99:139-142(2000).
DR  EMBL; AB032467; BAB19807.1; -.
DR  HSP; Q58083; IHV8.
DR  InterPro; IPR001410; DEAD.

```

```

DR  InterPro; IPR000629; DEAD_box.
DR  InterPro; IPR000362; Fumarate_lyase.
DR  InterPro; IPR001650; Helicase_C.
DR  InterPro; IPR000685; Rubisco_large.
DR  Pfam; PF00270; DEAD; 1.
DR  Pfam; PF00271; helicase_C; 1.
DR  SMART; SM00487; DEXDC; 1.
DR  SMART; SM00490; HELIC; 1.
DR  PROSITE; PS00039; DEAD ATP HELICASE; 1.
DR  PROSITE; PS00163; FUMARATE_LYASES; UNKNOWN_1.
DR  PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
DR  ATP-binding; Helicase.
SQ  SEQUENCE 645 AA; 70072 MW; 4C94536A9BF21DA CRC64;

Alignment Scores:
Pred. No.: 4,63e-137 Length: 645
Score: 1865.00 Matches: 369
Percent Similarity: 68.80% Conservative: 94
Best Local Similarity: 55.41% Mismatches: 135
Query Match: 47.17% Indels: 84
DB: 13 Gaps: 18

US-09-714-865-15 (1-2172) x Q9DEG3 (1-645)
Qy      109 ACTCAGCTTCATCATCAGAAATGATGATGACCTTCTCGAAGATCATTTGATGAAA 168
Db      14 Thrcyalaleuthrserapthrserthrnglyglthrclnglyasp----- 29
Qy      169 AGTGAATTTGCTCTGGCGGAAATTTGGAACACAGATGCTGTGATGATTAAGCA 228
Db      30 -----PheTPAsnserasphehlglylu----- 37
Qy      229 GATATATCATCCACATGSGTGTGTTGGAGTTGGAAGTTTGGAAACAGAGTTT 288
Db      38 -----Pheglndlygllyargly---GlyArgglyArgglyArgglylyPhe 52
Qy      289 TCACACAGAGAGTTT-----GAAGATGTGATGATGCTGTGTTCTGAGAGAGTCT 319
Db      53 LysAsnserPhehsergllyalgluaspgllyAlaAsnaspoly----- 67
Qy      340 AGTAATGACTGCGAGATATATCCACACGAAACAGAGGTTTTCACAGAGCGGCTAT 399
Db      68 AspGlyAspGlyGlnAsnAsnThrAlaGluArgSerSerPheSerArgArgglyly--- 86
Qy      400 CGAGATGAAATATTCAGACAGCTTCAGAGGCATACAGAAAGAGGTAGTATTC 459
Db      87 -----ArgglyArgglylys--- 91
Qy      460 CGAGTTGCCGTGAGAGATTTGGTCTAGAGAGTCCAAATATGACTTAGACCCAGAGAA 519
Db      92 -----GlyPheglly---ArgThrAspHisSeraspPheaspAlaAngln 105
Qy      520 TGTATCAGCGACATGCTGCTTTTGGTCTAGAAACCAAGTATTAAGTGCACAGT 579
Db      106 GluaspGlyAspAsnlysaAsnvalserArglyArgArg-----ArgglyArggly 122
Qy      580 AATGTGATTAATCTTCAAGACGAAGTGGCAGTGGAGAGGAAGAGTGGTTTCAAGGT 639
Db      123 GlyGlyglnglly---GlyArgserglyPheglly-----GlyglylyArggly 138
Qy      640 TTAAATGAGAAATATTAACAGGCTCTGGAAGAAATTTCTTGAAGTACAGAA----- 690
Db      139 LysaspGlu-----LysSerleuLeuylsAlaArgglylyrPro 151
Qy      691 GCAGAGAGAGAGAAAGTATGATCTCAAGACCAAAAGTGCATCATACCCCTCTCT 750
Db      152 AlagluAsnAlaspheleasparg-----ProArgValThrlyrValPropPro 169
Qy      751 CCACCTGAGATGAGAGATCCATCTTTCGACATTAACAGAGCGGCTTAACCTCGACAAA 810
Db      170 ProSerGluaspilndlyserillePheSerhstryGluSerllyAlaAsnAspLys 189
Qy      811 TACGACACTATTTCTGTGAGAGTCTGACATGATGACACACAGCAATTCGACTTT 870

```

Db 190 TyrAspGluIleValAspValSerCylThirsnProProProlaIleMetThrPhe 209

QY 871 GAAAGAGCTATCTCTGTCCAGACACTGATTAACAACATGCTTAAGCTTAACTAG 930

Db 210 AspGluIaIaIeucyGsluSerIeuluyArghenValSerLySserGlyTyVallys 229

QY 931 CTTACTCCTGTGCAAAAATAACAGTATTCATCACTTGCAGAGACGAGATTGATGCT 990

Db 230 ProThrProValGlnlyshIeGlyIleProIleIleSerIaIaGlyArGAspLeuMetIa 249

QY 991 TGTCSTCAACAGGGCTCTGGGAACACGCGGCTTTCCCTCAACCAATTTTGGCCATATG 1050

Db 250 CysIaGlnThrGlySerGlyLySerThraIaIaPheLeuLeuProIleLeuGlnGlnleu 269

QY 1051 ATGATGATGAGAAATAACATGCGCAGTCGTTTAAAGGTGCGAGAAACGACGATGATATT 1110

Db 270 MetIaAspGlyIaIaIaIaIaSerCysPheSerIeulMetGlnIuProAspIaIeIle 289

QY 1111 GTACCACTCACTGAGAAATTGGTCAACCAGATTTATTTGGAAGCCAGAAAAATTTCTTT 1170

Db 290 ValIaIaProThrArGJuleuIleasnGlnIleTyLeuGlnIuIaIaArgLyPheAlaTyx 309

QY 1171 GGGACTTGCTGTAAGAAGCTTGTTATATATATGGGGAAACCCAGCTGGGACATTCAATTGCA 1230

Db 310 GlyThrCysValaAspProValValValTyIeGlyValaSerThrglyshIeGlnIleArg 329

QY 1231 CAAATAGTACAGAGCTGTAATATATATATATGCTCTACTCCTGGAAAGCTAGATGATCTA 1290

Db 330 AspLeuLeuArGJlyCysAsnValLeucySgIyThrProGlyIaArgLeuAspMetIle 349

QY 1291 GGCAAAGAAAAAGATTGCTCTCAAAAACAGATCAAAATACTAGTTTGGATGAAGCTGATGC 1350

Db 350 GlyArgGlyLySerValaGlyLeuThrLySerValaArgTyLeuValLeuAspGluIaIaAspArg 369

QY 1351 ATGTTGGATATAGGGTTTGGTTCAGAAATGAAGAAGTTAATTTCTTGCCCGAGAAATGCCA 1410

Db 370 MetLeuAspMetGlyPheGlnIuProAspMetArgIeValaGlySerProGlyMetPro 389

QY 1411 TCAAAGGAACGCGCCAAACCTTATGTCAGTGCAACTTTCCAGAGAAATTCAAAG 1470

Db 390 SerLySgIuAsnArGJlnThrLeuMetPheSerIaIaThrPheProGlyAspIleGlnArg 409

QY 1471 TTGGCTCGACAGTTTAAAGTCAATATATCTGTTGTTCGCTGTGCAAGAGTGGGTGCA 1530

Db 410 LeuIaIaIaAspPheLeuLyThrAspTyLeuPheLeuIaIaValaGlyIleValaGlyIy 429

QY 1531 GCATGTAGAGATGTTCCAGACAGCGTTTCCAAAGTTGGCCAGCTTCAAAAAGAGAAAG 1590

Db 430 AlaCysSerArpValaGlnIuThrPheValaGlnValaThrLyPheAlaLyArGJln 449

QY 1591 CTGCTGGAATTTTCGCGAAACATAGGGGAGTAAAGAACTATAGGCTTTGTTGAACATAG 1650

Db 450 LeuIleAspLeuLeuLySerThrglySerGlnArghIuMetValaPheValaGlnThrLyS 469

QY 1651 AAAAAAGCAGATTTTACTGCAACTTTTCTTGTCAAGAAAAATAATCAACTACATCAAGTAC 1710

Db 470 ArgGlnIaAspPheIleIaIaThrIleLeucySglnIuLySphenroThrThrSerIle 489

QY 1711 CATGTGATCGGGAACGAGAGAGCGGAGCAAGCTCTTGGAGATTTTGCCTTTGGAAG 1770

Db 490 HisGlyAspArgGlnIuThrPglIaArgGlnIuIaIeGlnIuAspPheArGSerCylLyS 509

QY 1771 TGCCAGATCTTGTTGCTACTTCAGTACGTGCGCAGAGGGCGGATATGAAAGATGTGCA 1830

Db 510 CysSerValaIeulValaIaThrSerValaGlyIaIaArgLyLeuAspIleAspAspAlaGln 529

QY 1831 CATGTTATCAATTTTGATCTTCTCTCAACATGATGAATATGTTTCATCGAATTTGGCGCT 1890

Db 530 PheValaIaIaAsnPheAspLeuProAsnMetIleAspGlyTyValaHisArgIleGlyArg 549

QY 1891 ACTGTCGTTGTGGGAATACGTGCGAGAGCAATTTCTTTTGTGATCTTGAATCGGATAC 1950

```

Db      550  ThrlGlyArCySgSLyAenThrlGlyArAlAValSerPheTyraaPrProGlyAlaAspSer 565
QY      1951  CATTAGACACACCCCTAGTAAAGATATGTCACAGATGCTCAACAGATGTTCTGCATGG 2010
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      570  GluLeuAlaArgSerLeuValThrLleuSerLysIaGlnGlnGluValProSerTyr 589
QY      2011  TTGGAGAAGATATGCTTTTAGTACATCAATTCCT---GGCTTCACTGGTAGTACAAGAGA 2067
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      590  LeuGlnGluPheAlaPheSerGlyHisSerProSerGlyPheAsnProProArgLysHis 609
QY      2068  AACGTGTTGCATAGTATTGATACCAAGAAAGGGCAAGAGCACTTGAAACACAGCTGGGCTTT 2127
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      610  -----PheAlaSerThrAspLeuArgLysGlyProGln-----GlyGlyPhe 623
QY      2128  -----TCTTCTTCAAGAGCTCCCAATCCAGATGATAGATGATCA 2166
Db      624  PheGlnAspGlySerValThrSerGlnLeuAlaIaGlnThrAlaAlaAspGlu 643
QY      2167  TGGGAT 2172
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      644  TrpGlu 645

RESULT 10
Q9DGR8 1D Q9DGR8 PRELIMINARY; PRT; 662 AA.
AC Q9DGR8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cyt.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20283536; PubMed=10821771;
RX Tsunekawa N., Naito M., Sakai Y., Nishida T., Noce T.;
RT "Isolation of chicken vasa homolog gene and tracing the origin of
RT primordial germ cells."
RL Development 127:2741-2750(2000).
DR EMBL; AB004836; BAB12337.1; -.
DR HSSP; Q58083; 1HW8.
DR InterPro; IPR002086; Aldenhyde_dehydr.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000685; RUBISCO_large.
DR Pfam; PF00270; DEAD. 1.
DR Pfam; PF00271; helicase_C. 1.
DR SMART; SMO0487; DEXDC; 1.
DR SMART; SMO0480; HELIC; 1.
DR PROSITE; PS00687; ALDENHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR PROSITE; PS00157; RUBISCO_LARGE; UNKNOWN_1.
KW ATP-binding; Helicase.
SQ SEQUENCE 662 AA; 73328 MW; BC244540133877FB CRC64;

Alignment Scores:
Pred. No.: 3,93e-129 Length: 662
Score: 1763.50 Matches: 377
Percent Similarity: 64.19% Conservative: 80
Best Local Similarity: 52.95% Mismatches: 164
Query Match: 44.60% Indels: 91
Db: 13 Gaps: 11

US-09-714-865-15 (1-2172) x Q9DGR8 (1-662)

QY 7 GATGAGAGTTGGAGAGCGAAGAAATCAACCCATATGTTCTTCTATGTTCCATATTTGAG 66
Db 2 GlnGluAspTrpAspThrGlnLeu-----Glu 10

```

QY 67 AAGATAGTATTCTGAGAAAATGAGACAATTTTAAAGACATCCAGCTTCATCATCA 126  
 Db 11 GlnGlnAlaAlaAlaAlaSerGlnGlnYArgSerGlnGlnGlnAlaTrpMetAlaAsnSer 30  
 QY 127 GAAATGATAGTACGACTTCTCGAAGATCATTTCAATGAAAAGATTGCTCTGGG 186  
 Db 31 G1YArgProAsnSerProSerLeuArg----- 39  
 QY 187 CGGAATTTGGAACAGAGATGCTGCGATGTAATAAGGAGATATACATCCACAATG 246  
 Db 40 -----PheSerSerArgPro-----SerSerProLeu 48  
 QY 247 GGTGCTTTTGAGATTGGAAGAGTTTGGAAACAGAGTTTTCMAACAGCAGATTGGA 306  
 Db 49 SerG1YArgProGlnYArgProAsnSerProPhePheG1YPheSerGlnAsnG1YSer 68  
 QY 307 GATGCTGATAGCTCTGCTTTCTGAGAGACTCTAGTATGATGATCGCAAGATTAATCCACA 366  
 Db 69 LeuG1YAlaAsnG1YLeu----- 75  
 QY 367 CGGAACAGAGGTTTCCAGAGA-----GGCGGCTATGAGATGGAATTAATTC 417  
 Db 76 ---AsnArgSerLeuProValGlnHisAspIleG1YG1YTrpSerG1YSerArgG1YSer 94  
 QY 418 -----GAGCTTCAGGCGCATACAGAGAGTGAGAGAGTAGTTTCCGA 462  
 Db 95 ValValArgGlnAsnArgGlnAspGlnProValThrArgPheG1YArgG1YArgSerSer 114  
 QY 463 GGTTCGCGTGAAGATTGGTCTAGAGATCCAAATATGATCTAGACCAAGCAATGT 522  
 Db 115 G1YSerArgAspPheGlnGlnArgAsnSerAlaAsnAspProGlnMetGlnAspGlnG1Y 134  
 QY 523 ATGAGAGCAGCTGGTGGCTTTTGGTTCAGAAAGACAGATTAATTAAGCAGATTAAT 582  
 Db 135 PheArgGlnValProG1YIlePheG1YGlnSerIleCysPheAsnSerGlnG1YArgAsn 154  
 QY 583 GGTGATCTCTCTCAAGC-----AGAAGTGGCAGTGGAGTGAACGA 624  
 Db 155 SerProLeuArgG1YSerProPheAlaProG1YG1YArgG1YAlaValG1YG1YProAla 174  
 QY 625 GGTGCTTACAAAGTTTAAATGAAGATTAACAGAGCTCTGGAAGAAATTTCTGGAG 684  
 Db 175 G1YAlaLeuYArgG1YArgSerGlnGlnIleAspSerG1Y----- 187  
 QY 685 TCAGAACAGAGAGAGAAAGTATGATCATCATCAAGCAAAAGTACATCATACC 744  
 Db 188 -----ArgG1YProG1YValThrIleValPro 196  
 QY 745 CCTGCTCCAGCTGAGATGAGACTCATCTTGGACATTAATGAGAGCATTAACCTTC 804  
 Db 197 ProProProProGlnAspGlnGlnSerIlePheAlaCysTrpGlnSerG1YIleAsnPhe 216  
 QY 805 GACAAATACGACACTATTTCTTGGAAAGTGTCTGACATGATGACACACAGCAATTTCTG 864  
 Db 217 AspArgTrpArgGlnCysAlaValGlnMetSerG1YLeuAspProAlaProLeuLeu 236  
 QY 865 ACTTTGAAGAGTATCTCTGCTGCACTGAAATGAACAATGCTTAAGCTGATAT 924  
 Db 237 AlaPheGlnGlnAlaAsnPheAlaGlnThrLeuArgG1YAsnIleSerIleValThrG1YTrp 256  
 QY 925 ACTAAGCTACTCTGTCGCAAAATATACAGTATCTTCACTACTTGAAGACAGATTTG 984  
 Db 257 SerIleSerLeuThrProValGlnIleHisSerIleProValIleGlnAlaG1YArgAspLeu 276  
 QY 985 ATGGCTTGTGCTCAACAGGCTCTGGAAAGATCGGCTTTTCTTCAACAATTTGGCT 1044  
 Db 277 MetSerCysAlaGlnThrG1YSerG1YLeuThrAlaAlaPheLeuProIleValAsp 296  
 QY 1045 CATATGATGATGATGATGATTAATCTGCAGATGCTTTTAAAGGTGACGAACAGATGT 1104  
 Db 297 ArgMetMetLeuAspG1YValThrAlaSer---PheProG1YSerGlnGlnAspProGlnCys 315  
 QY 1105 ATTATGTAGACCACTCGAGATTTGTTCAACAGATTTTATTTGGAAGCAGAAATTT 1164

Db 316 IleIleValAlaProThrArgGlnLeuIleAsnGlnIlePheLeuGlnAlaArgIlePhe 335  
 QY 1165 TCTTTTGGACCTTGTCTAGACCTGTTGTTATATATGGGGAAACCCAGCTGGACATTC 1224  
 Db 336 ValIleG1YThrCysIleArgProValIleIleTrpG1YIleThrGlnThrGlnHisSer 355  
 QY 1225 ATTGCAAAATAGTACAAAGCTGTATATATATATATGCTACTCTGGAAGACTGATGAT 1284  
 Db 356 IleArgGlnIleMetGlnIleCysAsnIleLeuCysAlaThrProG1YArgLeuLeuAsp 375  
 QY 1285 ATCATAGGCAAAAGAAAGATTGCTCAACAGATCAATTAATCTTACTTTGGATGAGCT 1344  
 Db 376 IleIleGlnIleG1YLeuIleSerLeuValGlnValIleTrpLeuValLeuAspGlnAla 395  
 QY 1345 GATCGATTTGGATATGAGCTTTTGGTCCAGAAATGAAGATTAATTTCTTGCCACAGA 1404  
 Db 396 AspArgMetLeuAspMetG1YPheG1YLeuAspMetLeuValLeuLeuIleSerTrpProGln 415  
 QY 1405 ATGCCATCAAGAGAACAGCGCCAAACCTTAATGTTTCAGTGAACATTTTCCAGAGAAAT 1464  
 Db 416 MetProSerLeuAspArgArgGlnThrIleLeuMetPheSerAlaThrPheProGlnGlnVal 435  
 QY 1465 CAAAGCTGGCTGCAGAGTTTAAATCAAAATATCTGTTTGTGCTGTGACAGATG 1524  
 Db 436 GlnArgLeuAlaG1YGlnPheLeuYSerThrAspTrpIlePheLeuValIleG1YAsnThr 455  
 QY 1525 GGTGAGCATGTAGAGATTTTCAAGACAGACCGTTTCCAGTTGGCCAGTTTCAAAAGA 1584  
 Db 456 CysG1YAlaCysSerAspValGlnGlnAsnIleLeuGlnValProArgLeuSerLeuAsp 475  
 QY 1585 GAAAGCTGTTGAATTTCTGCGAAACATAGGAGATGAAGAATATGCTCTTTGTGA 1644  
 Db 476 AspIleLeuIleGlnIleLeuGlnSerThrG1YGlnArgThrMetValPheValAsp 495  
 QY 1645 ACTAAGAAAAGACAGATTTTACTGCAACTTTTCTTGTCAAGAAAATATCACTACA 1704  
 Db 496 ThrIleLeuAlaAspTrpLeuAlaAlaPheLeuCysGlnGlnAsnLeuProSerThr 515  
 QY 1705 AGTATCATGATGATGCGGAGAGAGAGCGGAGCAAGCTTCTGGAGATTTTGGCTTT 1764  
 Db 516 SerIleHisG1YAspArgGlnGlnArgGlnValAlaLeuValArgAspPheArgSer 535  
 QY 1765 GGAAGTGGCCAGTCTTGTCTGCTACTCTAGTAGCTGCCAGAGGCTGATATTGGAAT 1824  
 Db 536 G1YLeuCysGlnIleLeuValAlaThrSerValAlaSerArgG1YLeuAspIleGlnAsn 555  
 QY 1825 GTGCAACATGTTATCAATTTTGAATCTTCTTACCATGATGATATGTTCAATG 1884  
 Db 556 ValGlnHisValIleAsnPheAspLeuProAsnThrIleGlnAspTrpValHisArgIle 575  
 QY 1885 GGGCGTACGTCGCTGTGGGAAATGCTGGCAGAGCAATTTCTTTTATCTGTAATG 1944  
 Db 576 G1YArgThrG1YArgCysGlnAsnThrG1YValAlaValSerPhePheAspArgGlnSer 595  
 QY 1945 GATTAACATTTGACAGAGCTCTAGTAAAGTATGACAGATGCTCAACAGATTTCTCT 2004  
 Db 596 AspG1YHisLeuValGlnSerLeuLeuValLeuSerArgThrGlnGln----- 612  
 QY 2005 GCATGTTGGAAGAAATGCTTTAGTACATCAATCTGCTTCAAGTGAATGACA 2064  
 Db 613 -----GlnPheGlnPheG1Y-----G1YArgMetAlaValGlnArg 624  
 QY 2065 GGAAGCTGTTGCATCAGTGTATACCAAGAGGC 2100  
 Db 625 ThrAsnIleValAlaSerThrTrpCysProG1YG1Y 636

## RESULT 11

Q9GNP1 PRELIMINARY; PRT; 770 AA.  
 AC Q9GNP1; 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)







Db	125	GLYGLYGLYGLYARGGLYSERGLYCYSRHELYSCYSGLYGLUGLUGLYHISEMERARG	144
QY	559	CCAGTATTAAAGTGGCAGCAGGTAAATGGTGATCTCTCAAGCAGAGTGGCATGGAGCT	618
Db	145	GLUCYSPROGLINGLYGLYGLYGLY-----ARGGLYSERGLYCYS	158
QY	619	GAAAGAGGGTGTTCAAAGGT---TTAAATGAAGAGTAATAACAGGCTCTGAAAGAGAT	675
Db	159	PHELYSCYSGLYGLUGLUGLYHISEMERARGGLUCYS-----	171
QY	676	TCTTGAAGTACAGAGCAGAGGAGAGANAAAGTAGTGTACTCTCAAGACCAAAAGTGACC	735
Db	172	PROARGENTHISRERGLYGLUGLYGLYULYSSERAPAG-----PROPTOLE---	188
QY	736	TACATACCCCTCTCCACCTGAGAGTAGAGACTTCACATCTTGACATTAATCAGACAGGC	795
Db	189	TYRILEPROPROPPOPPOPGLUASPILUVALGIUMETPHEALASERMETGLNARGGLY	208
QY	796	ATTAACCTTGCAAAATAGACACATTTCTTGTTGGAAGTGTCTGGACATGATGACACACCA	855
Db	209	ILEASNPEGLYLYSTRYASPALALLEPROVALGIVALSERGLYVALASNALAPROLYS	228
QY	856	GCAATTCGACTTTTGAAGAAGCTAATCTGTGTCGACACCTGAATTAACAATTCCTTAA	915
Db	229	SERILEPROTHRPHIEGLUVALALACLYNEUPROGLUTHRVALILEUALASNVALLYSERG	248
QY	916	GCTGGTTACTAAGCTTACTCTGTGTGCAAAATAACAGTATTCCTATCATCTTGACAGA	975
Db	249	ALASENTYRGLUARGPROTHRPROVALGINLYSTRYSERILEPROILLEASNALASP	268
QY	976	CGAATTTGATAGGCTGTGTGCTTCAAAACAGGGTCTGGGAAGACTGGGGCTTTTCTCTTACA	1035
Db	269	ARGASPLEUMETALACYSALAGLINTHRCLYSERGLYSTRALALAPHELEULEUPRO	288
QY	1036	ATTTTGGCTCATATGATCATGARGGAATTACTGCACTCGTTTAAAGATTGACAGAA	1095
Db	289	VALLEUTHRILYSLEULIETHRASNGLYNEUGINEERGLINPHESERGLYSGLINTHR	308
QY	1096	CCAGAGTGATATTATTGTAGACCAACTCGAATTTGTCACACAGATTTATTGGAAAGCC	1155
Db	309	PROARGALILEVALVALGLYPROTHRARGGLUULEUILETRYGINLLEPHELEUGLUALA	328
QY	1156	AGAAAATTTTCTTTGGGACTTGTGTAAAGACTGTTGTTATATATGGGGGAAACCCAGCTG	1215
Db	329	ARGLYSPHESERARGLYTHRVALVALARGPROVALVALATYRGLYGLYTHRSERMET	348
QY	1216	GGACCTTCAATTGACGAAATGTCACAGGCTGAATATATATATGTCACCTCCGGAAGA	1275
Db	349	ASNHSISGLINLIEARGASPLEULEUNGILYCYSHSISLLEULILEALATHRPROGLYARG	368
QY	1276	CTGATGTGATCATAGGCAAGAAAGAAATGGTGTCAAAACAGATCAACTTACTGATTTTG	1335
Db	369	LEUWETASRPHELLEASNARGGLYNEUVALGILYLEUASRPHLSVALGILUPHEVALILLEU	388
QY	1336	GATGAAGCTGATCCGATGTTGGATATGGGTTTTGGTCCAGAAATGAAGAAGTTAATTTCT	1395
Db	389	ASRGLUALAASPARGMETLEUASPERCGLYHIEGILUTHRGILULEARGLYLEUALASER	408
QY	1396	TGCCCAGAGAAATGCCATCAAGAAGACAGGCCAAACCCCTATGTTCAAGTCACTTTTCCA	1455
Db	409	SERPROGLYMETPROSERLYSERASPRIGHISITHRLEUWETPHESERVALATHRPHERO	428
QY	1456	GAGGAAATTCCAAAGGTGGCTGCAGAGATTTTAAAGTCAAAATTTATCTTTGGTGGCTTT	1515
Db	429	ASRGLUILEGINARGLEUALAHIEASRPHIEUARGGLUASRPHELEUPHELEUTHRVAL	448
QY	1516	GGACAGTGGGTGAGCATGTGAGATGTTCCAGACAGACCGTGTCCCAAGTTGGCCAGTTC	1575
Db	449	GLYARGVALGILYGLYALACYSSTRASRYVALTHRGINSERTILLEGINVALASRGLINASP	468
QY	1576	TCAAAAAGAGAAAGCTGCTGGAATTTCTGGCAAAACATAGGGGAT-----GAAAGAACT	1629
Db	469	ASRYLSARGALALYSLEULEUGLILEUULESERPHEVALALAGLUTHRARGSERARGTHR	488

QY	1630	ATGGCTCTTGTGAAACATCAAGAAAAGACATTTTCTGCAACTTTTCTTGTGCAAGA	1669
Db	489	LeuvalPhevalGluThrLysArgGlyAlaAspPheLeuAlaCysMetLeuSerGlnGlu	508
QY	1690	AAATATCAATCAATCAATATCATGTGATCGGAAACAGAGACGGAGCAAGCTCTT	1749
Db	509	GlyCysPProThrThrSerIleHisGlyAspArgLeuGlnGlnIleValSerGluAlaLeu	528
QY	1750	GGAATTTTGGCTTTGCAAAAGTCCCAAGTTTGTGTGCTACTTACAGTACGCTCCAGAGG	1809
Db	529	ArgAspPheLysSerAlaValAlaCysProIleLeuIleAlaThrSerValAlaAlaArgGly	548
QY	1810	CTGGATATATGAAAAATGTGCACACATGTTATCAATTTGATCTCTCTTACCATGATGAA	1869
Db	549	LeuAspIlePProLysValGlnHisValIleAsnTyrAspMetProLysGluIleAspGlu	568
QY	1870	TATGTTCAATCAATGGGGCGTACTGTCTGTGTGGAAATCTGGCAGACAAATTTCTTT	1929
Db	569	TyrValHisArgGlnIleGlyArgThrGlyArgGlySerLeuGlnIleAlaThrThrPhe	588
QY	1930	TTTGATCTCTGAATCGGATACCATTTTACACAGCCTTACGTAAAAGTATTCACAGATGCT	1989
Db	589	TyrAspAsnAsnLysAspGlyGluIleValAlaArgSerLeuValLysIleLeuSerGluAla	608
QY	1990	CAACAGAGTGTCTCCGTCATGTGTGGAAGAAATTCCTTTGATACATCATCTCGGCTTC	2049
Db	609	GlnGlnGlnValAlaProGlyTyrPheGlnGlnGlnCysAlaGlnSerAlaVal-----	624
QY	2050	AGTGGTAGTACAGAGGAAACGTG-----TTTGATCACTGTGTATACACAGAG-----	2097
Db	625	---GlySerSerPheGlyLysGlnGlyGlyPheGlyGlyArgAspPheArgLysArgGly	643
QY	2098	-----GGCAAGACACTTTGAACACAGCT	2121
Db	644	GlyArgAsnAspThrGlyPheGlyLysSerLysGlyGlyPheGlySerArgSerThrAla	663
QY	2122	-----GGGTTTCTCTTCACAGAGCTCCCAATCCAGTAGAT	2157
Db	664	AspCysAspTyrAsnAspGlyGlyGlyPheGlyAsnThrAlaAlaValSerLysAspAsp	683
QY	2158	GATGAGTCATGGAT	2172
Db	684	AspAspSerTyrAsp	688
RESULT 13			
QY	096069	PRELIMINARY; PRT; 659 AA.	
AC	096069;		
DT	01-MAY-1999 (TREMBlrel. 10; Created)		
DT	01-MAY-1999 (TREMBlrel. 10; Last sequence update)		
DT	01-MAR-2002 (TREMBlrel. 20; Last annotation update)		
DE	DEAD-box protein.		
GN	CIDEADIA.		
OS	Ciona intestinalis.		
OC	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;		
OC	Phlebobranchia; Clonidae; Ciona.		
OX	NCBI_TaxID=7719;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=OVARY;		
RA	Takamura K., Fujimura M.;		
RT	"Expression Patterns of Ascidian DEAD-Box Genes.";		
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.		
CC	-I- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.		
DR	EMBL; AB016604; BAA56711.1; -.		
DR	HSSP; Q58083; 1HV8.		
DR	InterPro; IPR001410; DEAD.		
DR	InterPro; IPR000629; DEAD box.		
DR	InterPro; IPR001650; Helicase C.		
DR	InterPro; IPR001878; Znf_CCHC.		
DR	Pfam; PF00270; DEAD. 1.		
DR	Pfam; PF00271; helicase C/ 1.		



QY 2092 AGAAG-----GGCAGAGCACT 2109  
 Db 613 ArgGluArgGlyGluGluGlyLeuThrGlyPheGlySerMetGlyGlySerGlyAla 632  
 QY 2110 TTGACACAGCT-----GGGTTTCTTTTCACAGAGCTCCC 2145  
 Db 633 ArgSerThrAlaAspTyrAspTyrAsnAspGlyGlyPheAspCysValGlyAlaAsn 652  
 QY 2146 AATCCAGTAGATGATGATGATGATGAT 2172  
 Db 653 Asp-----AspAspAspSerTyrAsp 659  
 RESULT 14  
 ID 096068 PRELIMINARY; PRT; 669 AA.  
 AC 096068;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAY-2002 (TReMBLrel. 20, Last annotation update)  
 DE DEAD-box protein.  
 GN CIDEAD1B.  
 OS Clona intestinalis.  
 OC Eukaryote; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 OC Phlebobranchia; Clonidae; Clona.  
 NC NCB1\_TaxID=7719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=OVARY;  
 RX MEDLINE=20130953; PubMed=10664149;  
 RA Fujimura M., Takamura K.;  
 RT "Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific  
 RT expression in the germ cells and its mRNA localization in the  
 RT posterior-most blastomeres in early embryo."  
 RL Dev. Genes Evol. 210:64-72 (2000).  
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.  
 DR EMBL; AB016603; BAA36710.1; -;  
 DR HSSP; Q58083; IHW8.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR000629; DEAD box.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR001878; Znf\_CCHC.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF00098; zf\_CCHC; 3.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR SMART; SM00343; Znf\_C2HC; 3.  
 DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
 DR ATP-binding; Helicase; RNA-binding.  
 KW SEQUENCE 669 AA; 72785 MW; 1F5C5BA546A16D8 CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2.83e-110 Length: 669  
 Score: 1522.00 Matches: 339  
 Percent Similarity: 61.32% Conservative: 89  
 Best Local Similarity: 48.57% Mismatches: 188  
 Query Match: 38.49% Indels: 82  
 DB: 5 Gaps: 18  
 US-09-714-865-15 (1-2172) x 096068 (1-669)  
 QY 196 GGAACAGAGATGCTGTGATGATTAAGACAGATATATACATCCACATG----- 246  
 Db 15 GlAsnThrAspPheArgGlyAlaSerSerPheAspAsnProTyrSerLysAspAsp 34  
 QY 247 GGTGGCTTTTGAAGTGAAGAGTTTGA---AACAGAGTTTTCACACAGCAGG--- 300  
 Db 35 GlyGlyPheGlySerGly-----PheGlyAspSerArgGlyGlyAlaArgSerLysGly 52  
 QY 301 ---TTTGAAGATGCTGTGATGCTGTGCTTTCGAGAGAGTCTAGTATGACTGGAGAT 357  
 Db 53 CysPheLysCysGlyGluGly-----HisMetSerArgGluCysProGln 68

QY 358 AATCCAAACAGCAAGAGGCTTTTCCAGAGAGCGGCTATGAGATGGAATTAATTCA 417  
 Db 69 AsnThrGlySerGlyPheGlyAspSer---ArgGlyGlyAlaArg----- 82  
 QY 418 GAAGCTTCAGGCCCATACAGAAGAGTGAAGAGTGTGTTCCGAGGTGCCGAGAGA 477  
 Db 83 ---SerLysGlyCysPheLysCysGlyGluGluGlyHisMet----- 95  
 QY 478 TTTGTCTAGGAAGTCCAAATTAATGACTTAGACCACAGACGAATGATGACCGCAGTGT 537  
 Db 96 -----SerArgGlyAspProGlnAsnThrGly 104  
 QY 538 GGCCCTTTTGGTCTTGAAGACCAAGTATTAAGTGCAACAGTATGCTGATCTCAA 597  
 Db 105 SerGlyPheGlyAspSerArg-----GlyGlyAlaArgSerLys 117  
 QY 598 AGCAAGAGTGGCAGTGAAGTGAACGAGGTGTTACAAAGTTTAAATGAAGATATA 657  
 Db 118 GlyCysPheLysCysGlyGluGlu-----GlyHisMetSerArgGlyCysProGlnAsn 135  
 QY 658 ACAGGCTCTGAAAAGATCTTGGAAGTCAAGACGAAGAGAGGAA----- 705  
 Db 136 ThrGlySerGlyAspArgHisSerAsnAlaTyrPheLysGlyGlyAspHisAlaAlaGln 155  
 QY 706 -----AGTAGTACTCAAGACCAAA---GTGACCTCATATACCCCTCT 750  
 Db 156 GluTyrHisLysAlaGlyAspGlyAspLysProArgProGluProLysThrProProPro 175  
 QY 751 CCACCTGAGATGAGAGCTCATCTTTGACACATTATAGACAGGCAATTAACCTTGACAAA 810  
 Db 176 ProProGluAspGluValGluMetPheAlaSerMetGlnArgGlyLeuAsnProGlyLys 195  
 QY 811 TAGGACACTTCTTGTGGAAGTGTGCGACATGATGACACACACAGCAATTCGACTTT 870  
 Db 196 TyrAspAlaIleProValGluValSerGlyLeuAsnAlaProLysCysIleSerThrPhe 215  
 QY 871 GAAGAGCTAATCTCTGTCAAGACATGAATTAACACATTCCTAAGCTGTATCTAAG 930  
 Db 216 GluMetAlaAsnLeuGluGlnThrIleLeuValAsnValGlnLysAlaGlyTyrAspArg 235  
 QY 931 CTTACTCTGTGCAAAAATACAGTATTCTATCAATCTTCAGAGACAGATTGATGCT 990  
 Db 236 ProThrProValGlnLysTyrSerIleProIleIleAsnAlaAspArgAspLeuMetAla 255  
 QY 991 TGTGCTCAACAGAGCTGTGGAAGCTGCGGCTTTCTCTACCAATTTGGCTCATATG 1050  
 Db 256 CysAlaGlnThrLysSerGlyLysThrAlaAlaPheLeuLeuProValLeuThrLysLeu 275  
 QY 1051 ATCCATGATGAATTAATCTGCCAGTCTGTTTAAAGAGTTCGAGAGACCAAGATGATATT 1110  
 Db 276 ValGluSerGlyValLysSerSerGluPheSerGluLysThrProGlnAlaIleIle 295  
 QY 1111 GTAGCAACCACTCGAAGATTTGTCAACACAGATTATTGGAGCCGCAAAATTTCTTTT 1170  
 Db 296 IleGlyProThrArgGluLeuValAsnGlnIlePheLeuGluAlaArgLysPheSerArg 315  
 QY 1171 GGAGCTGTGTAGAGCTGTGTTATATATATGAGGGAACCAAGCTGGACACTTCAATTGCA 1230  
 Db 316 SerThrIleHisProValValIleValIleValIleThrSerValGlyTyrGlnIleArg 335  
 QY 1231 CAATAGTACAAAGCTGTATATATATATATGCTACTCTCGAAGACATGATGATATCATATA 1290  
 Db 336 AlaValGlnLysGlyCysAspValIleIleIleIleThrProGlyLysLeuMetAlaPheIle 355  
 QY 1291 GGCAGAAAAAAGATGTGCTCAAAACAGATCAAAATTAATCTGATTTTGAAGTGAATGCGC 1350  
 Db 356 AsnArgGlyLeuIleGlyLeuGluAsnValArgPheIleLeuLeuAspGluAlaAspArg 375  
 QY 1351 ATGTGGATATGGGTTTGTGTCAGAAATGAAGAAGTATTTCTTCCAGAGAAATGCCA 1410  
 Db 376 MetLeuAspMetGlyPheGluSerGluIleArgLysLeuValSerLeuProGlyMetPro 395  
 QY 1411 TCAAGAGCAAGCGCAAAACCTTATGTTCAAGTCAAGCTTTTCAAGAGAAATTCAAAG 1470

```

Db      396 GlnlyAsnAsnAargHisThrLeuMetPheSerIaThrPheProAspGluIleGlnIlye 415
Qy      1471 TTGGCTCCAGAGTTTAAATCAATTAATGCTTTGCTGTTGTCAGCAAGTGGGCGA 1530
Db      416 LeuLahIsapPheLeuAargGluAspPheLeuPheLeuThrValGlyAargValGlyGly 435
Qy      1531 GCATGTAGAGATGTCAGACAGCCGCTTCCCAATGGCCAGTCTCAAAAAGAGAAAAG 1590
Db      436 AlacysSerAspValIthrGlnThrValIleSerValGluTrpIlyAspIlyAspSerIlye 455
Qy      1591 CTGCTGAATTCGCGAAACATAGGGGAT-----GAAAGACTATGCTTTGTTGAA 1644
Db      456 LeuLeuGluLeuIleAlaAspValIaSerGluThrIlySerAspGluThrLeuValPheValGlu 475
Qy      1645 ACTAGAAAAAGAGATTTTACTGCAACTCTTGTGTCAGAAAAAATCAACTACA 1704
Db      476 ThrIlyAspGlyAlaAspPheLeuAlaCysValIleuCysGlnGluAspPheProIthrThr 495
Qy      1705 AGTATCCATGCGTGCAGGAACAGAGAGCGGAGCAAGCTCTTGAGATTTTCGCTTT 1764
Db      496 SerIleHisGlyAspAargLeuGlnIaAspAargGluGlnAlaLeuAargPheIlyleu 515
Qy      1765 GAAAGTCCCGCAGTTCTTTGTTGCTTCACTTCACTGCGCAGAGGCGCTGAGATTTGAAAT 1824
Db      516 AlavalCysProIleLeuValAlaIaThrSerValAlaIaAargGlyLeuAspIleProIly 535
Qy      1825 GTGCAACATGTTATCAATTTTGATCTCTTCACTCACTGATGTAATATGTCATCGAATT 1884
Db      536 ValIleHisValIleAsnIlyAspMetProAargGluIleAspGluIlyValHisAargIle 555
Qy      1885 GGGGCTACTGCTGTTGGGAATATCGAGACCAATTTCTTTTGTGATCTTGAATCG 1944
Db      556 GlyIargThrGlyAargCysGlyAsnLeuGlyIargAlaIthrThrPhePheAsnIlyeIys 575
Qy      1945 GATACCATTTAGACAGACCTCTAGTAAAGATTTGACAGATGCTCAACAGATGTTCT 2004
Db      576 AspAlaAsnLeuAlaAargSerLeuValIlyIleLeuSerGluAlaGlnGlnIValPro 595
Qy      2005 GCATGTTTGGAAATTTGCTTGTAGTACATACATCTCTGCTGCTCAGTGGAGTACAGA 2064
Db      596 GlyIlyPheGlyGlyCysAlaGluSerAlaValGlySerAsnIlye---GlyAlaGluIys 614
Qy      2065 GAAAGCTGTTTGCATCAGTTGATACAGAAAG----- 2097
Db      615 GlyIarg---PheGlyGlyIlyAspLeuAargGluAargGlyIleThrGlyPhe 633
Qy      2098 -----GGCAAGACACTTTGAACACAGCT----- 2121
Db      634 GlySerMetGlyGlyIlySerGlyAlaAargSerThrAlaAspIlyIlyAspIly 653
Qy      2122 ---GGGTTTCTTCTTCCACGAGCTCCCAATCCAGTAGATGATGATGATGATGATGAT 2172
Db      654 GlyIlyPheAspCysValGlyValAsnAsp-----AspAspAspSerTrpAsp 669

```

## RESULT 15

```

08QGD0 ID 08QGD0 PRELIMINARY: PRT: 394 AA.
AC 08QGD0:
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DE DEAD box RNA helicase Vasa.
OS Pantodon buchholzi (Butterflyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossiformes; Pantodontidae; Pantodon.
NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21906530; Pubmed=11909530;
RA Knaut H., Steinheisser H., Schwarz H., Nusslein-Volhard C.;
"An Evolutionary Conserved Region in the vasa 3'UTR Targets RNA

```

```

RT Translation to the Germ Cells in the Zebrafish."
RL Curr. Biol. 12:454-466 (2002).
DR EMBL: AF479823; AAL87142.1; -.
KW Helicase.
SQ SEQUENCE 394 AA; 43662 MW; 78F9ADFLA40A324 CRC64;
Alignment Scores:
Pred. No.: 2,19e-108 Length: 394
Score: 1497.00 Matches: 288
Percent Similarity: 84.75% Conservative: 51
Best Local Similarity: 72.00% Mismatches: 51
Query Match: 37.86% Indels: 10
DB: Gaps: 3
US-09-714-865-15 (1-2172) x 08QGD0 (1-394)
Qy      985 ATGGCTTGTGCTCAAAACAGGCTCTGGGAAGACTCGCGCTTTCTCTACCAATTTGGCT 1044
Db      1 MetAlaCysAlaGlnThrGlySerGlyIleThrAlaAlaPheLeuLeuProIleLeuGln 20
Qy      1045 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104
Db      21 GlnLeuMetThrSerGlyValAlaIaSerSerArgPheSerGluValGlnGluProGluAla 40
Qy      1105 ATTATGTAGCACCAACTCGAGAATTGTCACAGATTTATTTGGAAGCCAGAAATTT 1164
Db      41 IleValValAlaProThrAargGluLeuIleAsnGlnIleIlyThrLeuGlnAlaAargIlyPhe 60
Qy      1165 TCTTTTGGGACTGTGTAAGACTGTTGTTATATATGAGGGGAACCCAGCTGGGACATTC 1224
Db      61 AlaPheGlyThrCysValAargProValValIlyIleGlyIlyThrMetIlyThr 80
Qy      1225 ATTCGAATATGACAGAGCTGTAATATATATATATATATATATATATATATATATATATAT 1284
Db      81 IleArgIleValIleuIlyGlyCysAsnIleLeuCysGlyIlyThrProGlyAargLeuAsp 100
Qy      1285 ATCATAGCAAGAAAGATTTGCTTCAACAGATCAAAATATAGTTTGGATGAAGCT 1344
Db      101 IleIleGlyIargIlyIysValGlyLeuAsnIlyLeuAargIlyLeuValIleAspGluAla 120
Qy      1345 GATTCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404
Db      121 AspArgMetLeuAspMetGlyPheGluProAspMetArgIleLeuValAlaIaSerProGly 140
Qy      1405 ATGCAATCAAGGAAGAGCCCAACCCCTATGTCAGTCACTTTCAGAGGAATTT 1464
Db      141 MetProProIlyGlnGluAargIleIthrLeuMetPheSerAlaIlyIlyProGluAspIle 160
Qy      1465 CAAGGTGGCTCGAGATTTTAAAGTCAATTAATATGTTGCTGTTGGAAGAAGT 1524
Db      161 GlnAargLeuAlaIaAspPheLeuIlyAargAspIlyLeuPheLeuValIleGlyGlnVal 180
Qy      1525 GGTGAGCATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
Db      181 GlyIlyAlaCysAsnAspValIleGlnIleValIleValIleValIleSerGlnIlySerIlyAarg 200
Qy      1585 GAAAGCTGTTGAATTTGCGGAAACATAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1644
Db      201 GlnIleLeuValGluIleLeuAargIthrThrGlyAsnIlyAargIlyIlePheValGlu 220
Qy      1645 ACTAGAAAAAGAGATTTTACTGCAACTTTTCTTTGTCAGAAAAAATATCAACTACA 1704
Db      221 ThrIlyAspGlyAlaAspPheIleAlaIthrPheLeuIlyGlnIleSerThrThr 240
Qy      1705 AGTATCCATGCTGATCGGAAACAGAGAGCGGAGCAAGCTCTTGAGATTTTCGCTTT 1764
Db      241 SerIleHisGlyAspAargGlnIaAargGluAargGluIleuGlnIleuGlnIlyAspPheAargSer 260
Qy      1765 GAAAGTCCCGCAGTTCTTTGTTGCTTCACTTCACTGCGCAGAGGCGCTGAGATTTGAAAT 1824
Db      261 GlyIlyCysProValIleuValAlaIthrSerValAlaIaAargGlyLeuAspIleGluHis 280
Qy      1825 GTGCAACATGTTATCAATTTTGATCTTCTTACATCAATGATGATGATGATGATGATGATGAT 1884

```

```
Db      281 ValGlnHisValValAsnPheAspLeuProSerThrIleAspGluTyrValHisArgIle 300
QY      1885 GGGCGTACTGGTGGTGGGAACTACTGGCAGAGCAATTCTTTTGTGATCTTGAATCG 1944
Db      301 GlyArgThrGlyArgCysGlyAsnThrGlyLysAlaIleSerPhePheAspProGluAla 320
QY      1945 GATTAACCATTTAGACAGCCTCTAGTAAAGTATTGACAGATGCTCAACAGATGTTCT 2004
Db      321 AspAlaProLeuAlaArgSerLeuValLysIleLeuSerGlyAlaGlnGlnGluValPro 340
QY      2005 GCATGGTTGAGAGAAATTGCCTTTAGTACATCAATTCTGCTTCAGTGGTAGTACA--- 2061
Db      341 LysTrpLeuGlnGluIleAlaPheSerAlaHis-----GlyThrThrAla 355
QY      2062 -----AGAGGAAACGTTGTCATGTCATGATACAGAGGAGGCAAGAGCACTTTG 2112
Db      356 PheAsnProArgGlyLysValPheAlaSerThrAspThrArgLys--GluGlyAlaPhe 374
QY      2113 AACACAGCTGGGTTTCTTCTTCACGAGCTCCCAATCCAGTAGATGATGATCATGGGAT 2172
Db      375 ProArgAlaGluProValGlnProAlaValGlnAsnProAspAlaAspGluGluTrpGlu 394
```

Search completed: June 10, 2003, 16:42:45  
Job time : 150 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:32:00 ; Search time 47.5 Seconds

(without alignments)  
8791.746 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 3954  
Sequence: 1 atgagggatgaagatcgga.....tagatgatgatgatggaat 2172

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

MODEL=frame+n2p.model -DEV=xlp  
-O=/cgp2.1/USPFO.spool/US09714865/runat\_05062003\_111758\_26034/abp\_query.fasta\_1.2311  
-DB=PIR\_73 -QFMT=fastan -SUFFIX=trp -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=Dicesum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09714865 @CGN 1.1 55 @runat\_05062003\_111758\_26034 -NCPU=6 -ICPU=3  
-NO\_MMAP -IARBOUDRY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MAIN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3323	84.0	713	2	T46407
2	3247.5	82.1	713	2	JC2534
3	2857.5	72.3	637	2	I49638
4	1959.5	49.6	700	2	I51235
5	1274	32.2	661	2	A58768
6	1229	31.1	662	1	I84741
7	1224.5	31.0	697	1	S13654
8	1197	30.3	660	2	A32378
9	1176	29.7	688	2	T48796
10	1126.5	28.5	604	2	S13653
11	1109	28.0	636	2	T43543
12	1108	28.0	617	2	S62003
13	1100	27.8	633	2	H84854
14	1095.5	27.7	646	2	T45677

15	1083.5	27.4	707	2	A48686	probable RNA helic
16	1069.5	27.0	603	2	T45671	ATP-dependent RNA
17	1026.5	26.0	604	2	C87818	protein gln-1 (imp
18	1026.5	26.0	604	2	T15132	ATP-dependent RNA
19	992.5	25.1	720	2	T15231	germline RNA helic
20	820.5	20.8	550	1	S14048	RNA helicase dbp2
21	815	20.6	546	1	S13757	RNA helicase DBP2
22	804	20.3	561	2	T22917	probable ATP-depen
23	797.5	20.2	619	2	T52137	ATP-dependent DEAD
24	796	20.1	575	1	S11485	RNA helicase - fru
25	790	20.0	501	2	B96593	probable ethylene-
26	777.5	19.7	614	1	I48385	RNA helicase TWZ2
27	777.5	19.7	614	2	UC1087	RNA helicase, ATP-
28	771.5	19.5	1156	2	T43326	germline RNA helic
29	771.5	19.5	1172	2	T32759	hypothetical prote
30	754	19.1	607	1	S42639	ATP-dependent RNA
31	754	19.1	650	2	S72367	ATP-dependent RNA
32	753	19.0	713	2	T48634	DRH1 DEAD box prot
33	741.5	18.8	1032	2	A57514	RNA helicase HELL1
34	713	18.0	748	2	H84913	probable ATP-depen
35	713	18.0	1166	2	H86341	hypothetical prote
36	712	18.0	446	2	E82835	ATP-dependent RNA
37	699	17.7	723	2	H84748	hypothetical prote
38	696.5	17.6	566	2	S53813	RNA helicase - sll
39	693.5	17.5	523	1	S30805	probable RNA helic
40	692.5	17.5	578	2	S67386	probable ATP-depen
41	688.5	17.4	537	2	A86444	probable RNA helic
42	685	17.3	454	2	E64816	probable ATP-depen
43	682.5	17.3	701	2	T15942	hypothetical prote
44	682	17.2	455	2	A10599	probable ATP-depen
45	681.5	17.2	455	2	C90738	probable ATP-depen

#### ALIGNMENTS

RESULT 1  
T46407  
probable RNA helicase protein DKFZp434B1122.1 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 17-Nov-2000  
R/Blum, H.; Baueersachs, S.; Mewes, H. W.; Gaessenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z23034  
A/Accession: T46407  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-635 <AAA>  
A/Cross-references: EMBL:AL137462; PIDD:CA870750.1  
A/Experimental source: adult testis; clone DKFZp434B1122  
A/Note: DKFZp434B1122.1

#### Alignment Scores:

Pred. No.:	2,1e-230	Length:	635
Score:	3323.00	Matches:	633
Percent Similarity:	99.84%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	1
Query Match:	84.04%	Indels:	0
DB:	2	Gaps:	0

US-09-714-865-15 (1-2172) x T46407 (1-635)

QY	268	AGCTTTGGAACGAGGTTTTC	CAACACGAGGTTGAAAGATGTGAT	CTGCTGTTTC	327
DB	1	SeppheglYasnrrgglYpneserAmserrArpghlunspglYaspserserglYpHe			20
QY	328	TGGAGAGAGCTCTGTATGATGCTGGAAGATTAATCAACACGAGGAGGTTTTC	CAAG		387
DB	21	TlprlgglusererAmsnabpCysglunspAmsproThrArgAmsnrrgglYpneserlYs			40
QY	388	AGAGGCGGCTATCGAGATGGAATTAATTC	GAAAGCTTCAGGCGCATACAGAGGTTGA		447

```

Db      41 ArglyglytyrArgAspGlyAsnAsnSerGluAsnSerGlyProtyrArgArgGlyGly 60
Qy      448 AGAGGTGTTTCGAGGTTGCGGAGGAGTTGGTCTAGGAGACTCCAAATTAAGACTTA 507
Db      61 ArglySerPheArgGlyCyArgGlyGlyPheGlyLeuGlySerProAsnAsnAspLeu 80
Qy      508 GACCACAGCATGATGATGACGACACTGGTGCCCTTTTGGTTTGAAGACAGCATATTA 567
Db      81 AspProAspGlyCysMetGlnArgThrGlyGlyLeuPheGlySerArgArgProValLeu 100
Qy      568 AGTGGCACAGGTAATGGTGATCTTCTCAAGCAGAAAGTGGCAGTGAAGTGAAGAGT 627
Db      101 SerGlyThrGlyAsnGlyAspThrSerGlnSerArgSerGlySerGlyArgGly 120
Qy      628 GGTACCAAGGTTAAATGAAGAAGTAAATACAGGCTCTGGAAGAAATCTTGGAAGTCA 687
Db      121 GlytyrArgGlyLeuAsnGluGluValIleThrGlySerGlyLysAsnSerTrpLysSer 140
Qy      688 GNAACAGAGAGAGAGAGAAAGTACTGATCTACAGGACCAAAAGTACCTACATACCCCT 747
Db      141 GluIleGluGlyGlyGlySerSerAspThrGlnGlyProLysValThrTyrIleProPro 160
Qy      748 CCTCCACTGAGATGAGAGACTCCATCTTGGCACTTATACAGACAGGATTAACCTTGAC 807
Db      161 ProProProGluAspGluAspSerIlePheAlaHisTyrGlnThrGlyIleAsnPheAsp 180
Qy      808 AAATACGACACTATTTCTGTGGAAGTGTCTGAGACATGATGACACAGCAAAATCTGACT 867
Db      181 LysTyrAspThrIleLeuValGluValSerGlyHisAspAlaProProAlaIleLeuThr 200
Qy      868 TTTGAGAGAGCTAAATCTCTGACAGACTGAAATACACATTTGCTAAAGCTGTTATACT 927
Db      201 PheGluGluIleAsnLeuCyGlnThrLeuAsnAsnAlaIleAlaLysAlaGlyTyrThr 220
Qy      928 AAGCTTACTCTGTGCAAAAATACAGTATTTCTTATCATCTTGGACAGACGAGATTGATG 987
Db      221 LysLeuThrProValGlnLysTyrSerIleProIleIleuAlaGlyArgAspLeuMet 240
Qy      988 GCTTGTCGTAAACAGGGGTCTGGGAGAGCTGGGCTTTTCCATACCAATTTGGCTCAT 1047
Db      241 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuAlaHis 260
Qy      1048 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
Db      261 MetMetHisAspGlyIleThrAlaSerArgPheLysGluLeuGlnGluProGluCysIle 280
Qy      1108 ATTGTAGCACAACCTCGAATTTGGTCAACAGATTTATTTGGAAGCCAGAAATTTTCT 1167
Db      281 IleValAlaProThrArgGluLeuValAsnGlnIleTyrLeuGluIleAlaArgLysPheSer 300
Qy      1168 TTTGGGACTTGTGTAGAGCTGTTGTTTATATGGGGGAGCCAGCTGGGACATTCAATT 1227
Db      301 PheGlyThrCysValArgAlaValAlaIleTyrGlyGlyThrGlnLeuGlnHisSerIle 320
Qy      1228 CGAACAATAGTACAAGGCTGTATATATATATATGCTACTCTCGAAGAGATGATGATATC 1287
Db      321 ArgGlnIleValGlnGlyCysAsnIleLeuCysAlaThrProGlyArgLeuMetAspIle 340
Qy      1288 ATAGGCAAAAGAAAGATTGTTCTCAACAGATCAAAATCTTAGTTTGGATGAAGCTGAT 1347
Db      341 IleGlyLysGluLysIleGlyLeuLysGlnIleLysTyrLeuValLeuAspGluAlaAsp 360
Qy      1348 GCGATGTTGGATATGCGGTTTGGTGCAGAAATGAAGAAGTAAATTTCTTGCCCGAGATG 1407
Db      361 ArgPheLeuAspMetGlyPheGlyProGluMetLysLysLeuIleSerCysProGlyMet 380
Qy      1408 CCATCAAAAGAAAGACGCGCAAAACCTTATGTTCAAGTCAACTTTTCCAGAGAAATTCAA 1467
Db      381 ProSerLysGluGlnArgGlnThrLeuMetPheSerAlaThrPheProGluGlnIleGln 400
Qy      1468 AGGTTGGCTGCAAGATTTTTAAAGTCAAAATATATCTGTTGTTGCTGTTGCAAGTGGGT 1527
Db      401 ArgLeuAlaIleAlaGluPheLeuLysSerAsnTyrLeuPheValAlaValGlyGlnValGly 420

```

```

Qy      1528 GGAGCATGTAGAGATGTTGACAGACAGCGTTTCTCAAGTTGGCCAGTTCTCAAAAGAGAA 1587
Db      421 GlyAlaCysArgAspValGlnGlnThrValLeuGlnValGlyGlnPheSerLysArgGlu 440
Qy      1588 AACCTGTTGAAATTTCTCGCAAAACATAGGGGATGAAAGAACTATGGTCTTTGTTAAACT 1647
Db      441 LysLeuValGluIleLeuArgAsnIleGlyAspGluArgThrMetValPheValGluThr 460
Qy      1648 AAGAAAAAGACATTTTACTGCACTTTTCTTGTTCAGAAAAAATATCAATCAACAGT 1707
Db      461 LysLysLysAlaAspPheIleAlaThrPheLeuCyGlnGlnLysIleSerThrSer 480
Qy      1708 ATCCATGCTGATGCGGACAGAGAGCGGAGCAAGCTCTTGGAGATTTTCCCTTTTGA 1767
Db      481 IleIleGlyAspArgGluGlnArgGluValArgGluGlnAlaLeuLysPheArgPheGly 500
Qy      1768 AAGTCCCACTTTGTTGCTACTTCACTGATGATGATGATGATGATGATGATGATGATG 1827
Db      501 LysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeuAspIleGluAsnVal 520
Qy      1828 CAACATGTTATCAATTTGATCTTCTTCTGACATGATGATGATGATGATGATGATGATG 1887
Db      521 GlnHisValIleAsnPheAspLeuProSerThrIleAspIleTyrValHisArgIleGly 540
Qy      1888 CGTACTGTCGTGTGTGGAAATACTGGCAGAGCAATTTCTTTTGTATCTTGAATCGAT 1947
Db      541 ArgThrGlyArgCyGlySerGlnThrGlyArgAlaIleSerPhePheAspLeuGluSerAsp 560
Qy      1948 AACCATTTAGCAGACGCTCTGATGAAAGTATGACAGATGCTCAACAGATGTTCTGCA 2007
Db      561 AsnHisLeuAlaGlnProLeuValLysValLeuThrAspAlaGlnGlnAspValProAla 580
Qy      2008 TGGTTGGAGAAATGCTTGTATGATACATTCCTGCGCTTCACTGATGATGATGATGATG 2067
Db      581 TrpLeuGluGluIleAlaPheSerThrTyrIleProGlyPheSerGlySerThrArgGly 600
Qy      2068 AACGTGTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2127
Db      601 AsnValPheAlaSerValAlaAspThrArgLysGlyLysSerThrLeuAsnThrAlaGlyPhe 620
Qy      2128 TCTTCTTACGAGCTCCCAATCCAGTATGATGATGATGATGATGATGATGATGAT 2172
Db      621 SerSerSerGlnAlaProAsnProValAspAspGluSerTrpAsp 635

```

## RESULT 2

```

JC2534
RVLG protein - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 20-May-1995 #sequence_revision 14-Jul-1995 #text_change 02-Feb-2001
C/Accession: JC2534
R:Komiyama, T., Tanigawa, Y.
Biochem. Biophys. Res. Commun. 207, 405-410, 1995
A/Title: Cloning of a gene of the DEAD box protein family which is specifically expressed
A/Reference number: JC2534; MIMD:95160706; PMID:7657296
A/Accession: JC2534
A/Molecule type: mRNA
A/Residues: 1-713 <COM>
A/Cross-references: GB:S75275; NID:g806463; PIRN:AAB3364.1; PID:g806464
C/Comment: This protein contains a conserved DEAD box.
C/Keywords: ATP; nucleotide binding; P-loop
F:317-324/Region: nucleotide-binding motif A (P-loop)
F:427-432/Region: nucleotide-binding motif B
F:431-434/Region: DEAD motif

```

## Alignment Scores:

```

Pred. No.: 5.65e-225 Length: 713
Score: 3247.50 Matches: 625
Percent Similarity: 90.55% Conservative: 36
Best Local Similarity: 80.62% Mismatches: 46
Query Match: 82.13% Indels: 23
DB: 2 Gaps: 6

```



Qy 2143 CCNATCCAGATGATGATCATGGGAT 2172  
 |||||  
 Db 704 ProbsnProValaAspAspGluSerTrpasp 713

RESULT 3  
 149638  
 Probable RNA helicase protein - mouse (fragment)  
 N:Alternate names: Drosophila vasa homolog  
 C:Species: Mus musculus (house mouse)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 02-Feb-2001  
 C/Accession: 149638  
 R:Fujiiwara, Y.; Komiya, T.; Kawabata, H.; Sato, M.; Fujimoto, H.; Furusawa, M.; Noce, T.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 12258-12262, 1994  
 A:Title: Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosop  
 A:Reference number: 149638; MUID:95083681; PMID:7991615  
 A:Accession: 149638  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-637 <RES>  
 A:Cross-references: GB:D14859; NID:g286074; PIDN:BAA03584.1; PID:g286075  
 C:Keywords: ATP; nucleotide-binding; P-loop  
 F:246-253/Region: nucleotide-binding motif A (P-loop)  
 F:356-361/Region: nucleotide-binding motif B  
 F:360-363/Region: DEAD motif

Alignment Scores:

Pred. No.:	5,25e-197	Length:	637
Score:	2857.50	Matches:	549
Percent Similarity:	91.39%	Conservative:	35
Best Local Similarity:	85.92%	Mismatches:	48
Query Match:	72.27%	Indels:	8
DB:	2	Gaps:	4

US-09-714-865-15 (1-2172) X 149638 (1-637)

Qy 253 TTGGAGCTGGAAAGAGTTTGGAAACAGAGCTTTTCAACACAGAGTTTGAAGATGT 312  
 |||||  
 Db 1 PheGlyArgGlyLysGlyPheGlyAsnArgGlyPheLeuAdeAsnLysPheGluGly 20

Qy 313 GATAGCTCTGGTTCTTGAGAGAGCTAGTAATGATCTGCCAAGTAATCCACACGGAGC 372  
 |||||  
 Db 21 AsperSerGlyPheTrpLysGluSerAsnAspCysGluAspAsnGlnThrArgSer 40

Qy 373 AGAGGAGTTTCCAGAGAGCGGCTATCGAGATGAAATATATAGAAAGTTTCAGGCGCA 432  
 |||||  
 Db 41 ArgGlyPheSerLysArgGlyGlyCysGlnAspLysAsnAspSerGluAlaSerGlyPro 60

Qy 433 TACGAAAGAGGTGAGAGAGTATGTTCCGAGTTGCCGTGAGAGATTTGCTAGGAAGT 492  
 |||||  
 Db 61 PheArgArgGlyLysArgGlySerPheArgGlyCysArgGlyLysPheGlyLeuGlyArg 80

Qy 493 CCAATATATGATGAGCCAGACGAAATGATGACGCGCATGCGTGGCTTTGGTTCT 552  
 |||||  
 Db 81 ProAsnSerGlySerAspGlnAspGlnGlyThrGlnCysGlyGlyPheLeuValLeu 100

Qy 553 AGAAGACAGATTAATAGTGGCAGAGTATGATGATGATCTTCAAGAGAGAGTGGAGT 612  
 |||||  
 Db 101 GlyLysProAlaAlaSerAspSerGlyAsnGlyAspThrTrpGlnSerArgSerGlySer 120

Qy 613 GGAAGTGAACGAGGTGTTCAAGAGTTTAATGAAGAAATATACAGAGCTTGAAG 672  
 |||||  
 Db 121 Gly-----ArgGlyGlyLysLysGlyLysLeuAsnGlnGluValValThrGlnLysSerGlyLys 138

Qy 673 AATTTCTGGAAGTGAAGAGCAGAGAGAGAGAGAAATGATGATCTCAAGACCAAAATG 732  
 |||||  
 Db 139 AsnSerTrpLysSerGluThrGlnGlyGlyGluSerSerAspSerGlnGlyProLysVal 158

Qy 733 ACCATCAATACCCCTCCATCGAGAGTGAAGATGAGATCTTCAATGACATTAACAGCA 792  
 |||||  
 Db 159 ThrTrpLysProProProProProGluAspGluAspSerLysPheAlaSerGlnThr 178

Qy 793 GGCATAAACCTTGAACAATACAGACATATTTCTTGTGAAAGTGTGACATGATGACCA 852  
 |||||

Db 179 GlyLysAsnPheAspLysTrpAspThrLysLeuValGluValSerGlyLysAspAlaPro 198  
 |||||

Qy 853 CCGAATCTGACTTTGAGAGAGCTAATCTGTGACAGACTGAATAACAATGCT 912  
 |||||

Db 199 ProAlaLysLeuThrPheGluGluAlaAsnLeuCysGlnThrLeuAsnAsnLysValArg 218  
 |||||

Qy 913 AAAGCTGTTATCTAAGCTTACTCTCTGCAAAAATAGATATTCATCTACTTGA 972  
 |||||

Db 219 LysAlaGlyLysThrLysLeuThrProValGlnLysTrpThrLeuValLeuAla 238  
 |||||

Qy 973 GAGCAGATTTGATGCTGTGCTCAACAGAGTCTGGAGACCTCGCTTTCTCTA 1032  
 |||||

Db 239 GlyArgAspLeuSerLysCysAlaGlnThrGlySerGlyLysThrAlaPheLeu 258  
 |||||

Qy 1033 CCAATTTGGCTATATGATGATGATGATGATGATGATGATGATGATGATGATG 1092  
 |||||

Db 259 ProLeuAlaAsnSerLysCysAlaGlnThrGlySerGlyLysThrAlaPheLeu 278  
 |||||

Qy 1093 GAACAGAGTGTATTTATGAGCACCACCACTGAGAAATGGTCAACAGATTTATGGA 1152  
 |||||

Db 279 GluProGluCysLysLysLeuValAlaProThrArgGluLeuLeuAsnGlnLysLeuGlu 298  
 |||||

Qy 1153 GCCAGAAAATTTCTTTGGAGCTTGTGATGAGCTGTGTTATATAGGGGACCCAG 1212  
 |||||

Db 299 AlaArgLysPheSerPheGlyThrCysValLysSerValValLysGlyThrGln 318  
 |||||

Qy 1213 CTGGGATTCATTCATGCAAAATGTAACAAGCTGTAATATATGATGATGATGATG 1272  
 |||||

Db 319 PheGlyLysSerValArgGlnLysValGlnGlyCysAsnLysLeuCysAlaThrProGly 338  
 |||||

Qy 1273 AGAGTATGATATCATAGAGCAAGAAAGATGCTGTCAACAGATCAAAATCTAGTT 1332  
 |||||

Db 339 ArgLeuMetAspLysLysGlyLysGlyLysGlyLysGlyLysGlyLysGlyLys 358  
 |||||

Qy 1333 TTGGATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1392  
 |||||

Db 359 LeuAspGlnAlaAspSerMetLeuAspMetGlyPheAlaProGluLysLysLysLys 378  
 |||||

Qy 1393 TCTTGGCCAGGATGCGATCAAGAGAAAGAGCGGCAACCTTATGTTGTCGACCTTT 1452  
 |||||

Db 379 SerCysProGlyMetProSerLysGluGlnLysGlnThrLeuLeuPheSerAlaThrPhe 398  
 |||||

Qy 1453 CCAAGAGAAATTCAAAGTTGGCTGCGAGATTTTAAAGTCAAAATATCTGTTGCT 1512  
 |||||

Db 399 ProGluGlnLysGlnArgLeuAlaGlyAspPheLeuLysSerThrLysLeuPheValAla 418  
 |||||

Qy 1513 GTTGGACAAAGTGGTGGAGCATGATGATGATGATGATGATGATGATGATGATG 1572  
 |||||

Db 419 ValGlyGlnValGlyGlyLysAlaCysArgAspValGlnGlnThrLysLeuGlnValGlyGln 438  
 |||||

Qy 1573 TTCTCAAAAAGAGAAAGCTCGTTGAATCTCTGCAAACTAGGGGATGAAAGCTATG 1632  
 |||||

Db 439 TyrGlnLysGlyLysSerLeuLeuArgPheTyrGluAsnLysLysGluArgThrMet 458  
 |||||

Qy 1633 GTCTTTGTTGAACATTAAGAAAAAGAGATTTTACTGCACTTTTCTTGTCAAGAAA 1692  
 |||||

Db 459 ValPheValGlnThrLysLysLysLysLysLysLysLysLysLysLysLysLys 478  
 |||||

Qy 1693 ATATCAATCAAGTATTCATGATGATGATGATGATGATGATGATGATGATGATG 1752  
 |||||

Db 479 LysSerSerThrSerLysLysLysLysLysLysLysLysLysLysLysLysLys 498  
 |||||

Qy 1753 GATTTTGGCTTGAAGTGGCCAGTTCTTGTGATGATGATGATGATGATGATGATG 1812  
 |||||

Db 499 AspPheArgCysGlyLysCysProValLeuValAlaThrSerValAlaAlaArgGlyLeu 518  
 |||||

Qy 1813 GATATGAAATATGCAACATGTTATCAATTTGATCTTCTTCAATGATGATAT 1872  
 |||||

Db 519 AspLysLysValGlnLysValLysLysLysLysLysLysLysLysLysLysLys 538  
 |||||

Qy 1873 GTTCATGAAATTTGGCGTACTGCTGTTGGGAAATACGAGAGCAATTTCTTTT 1932  
 |||||

Db 539 ValHisArgLysGlyArgThrGlyArgCysGlyAsnThrGlyArgAlaLysSerPhePhe 558  
 |||||

[illegible]

QY	352	-----GAAGTATATTCACCAACGAGAGGCTTTCCAAAGAGCGGCTATTCGA	402
DB	94	ValIheglYsphaSpGluIhrgaSpGlnIhrgIyIhPheProGlyIhArgGlyIhTyIhAsn	113
QY	403	GATGAAATATATTCAGAACTTCAGGGCCATACAGAAAGAGGTGGAGAGTTCCTCCGA	462
DB	114	GlyAnsgluAhpGlyIhGlnIhysProAhnAlaIhPheArg-----GlyIhArgGlyIhGlyIhPheArg	131
QY	463	-----GCTTGCCGTGGAGATTTGGCTCTAGGAATTCGA	495
DB	132	AsnGluAhnsgluGlnIhArgArgGlyIhPheGlyIhGluIhArgGlyIhGlyIhPheArgSerGluAhnGly	151
QY	496	AATATAGACTTAGCCCAAGCAAGATGTATGAGCGCACTGGTGCTTTTGGTTCAGAA	555
DB	152	GlnIhArgAhnPheAsp-----AsnIhArgGlyAhpPheGlyIhAsnSer	164
QY	556	-----AGACCAAGTATTAAGTGGCACAGCT-----AATGTGATTA	591
DB	165	GlyIhGluIhGluIhAspAhpGProIhArgSerTyrGlyIhArgGlyIhGlyIhPheAhnAsnSerAhpThr	184
QY	592	TCT-----CAAAGCAGAAGTGGCAGTGGAAAGTGAACGAGGTGTTACAAAGT	639
DB	185	GlyIhIhArgGlyIhArgArgGlyIhArgGlyIhGlyIhArgSerGlnIhGlyIhTyIhGlyGly	204
QY	640	TTAATATGAAGAAGTAAATACAGGCTCTGGAAAGAAATCTTGGAACTAGAAAGCAAGAA	699
DB	205	ArgAhnsgluGluVal-----GlyIhAlaIhSerGlyIhIhysSerGlnIhGluGly	220
QY	700	GGACAAAGTAGTAGTACTACAGAGCCAAAGATGACTACATACCCTCTCCACCTGAG	759
DB	221	AsnGlu---LysAhpGlyIhysProIhysValIhThryIleProProProProAhp	239
QY	760	GATGAGAGCTCCATCTTTGGACATTAATCAGACGACATAAATCTTGACAAATACGACT	819
DB	240	GlyIhIhAspAhnIlePheArgGlnIhTyIhSerGlyIhIhAsnPheAhpLysTyIhAspGlu	259
QY	820	ATTCTTGTGGAAGTGTCTGCACATGATGCACACCACCAATCTGCATCTTTGAAGAAGCT	879
DB	260	IleIhValIhAspValIhIhrgIhLysAhpValIhProIhAlaIleIhLeuIhIhPheGluGluVala	279
QY	880	AATCTGTGTGACACTGATTAACAAATGCTAAATGCTGTTATTAAGCTTACCTCT	939
DB	280	AsnIhIhysGluIhIhLeuAhpIhArgAhnValaIhArgAlaGlyIhValIhLysIhLeuIhPro	299
QY	940	GTCGCAAAATACAGTATTCCTATCATCTTGCAAGACGAAATTTGATGCTGTGCTCAA	999
DB	300	ValGlnIhysHISserIleProIleIleMetAlaGlyIhArgAhpIhMetAlaCyAlaGln	319
QY	1000	ACAGGCTGTGGAAAGACTCGGGCTTTCTCTCCACAAATTTGGCTCATATGATGATGAT	1059
DB	320	ThryIhSerGlyIhysIhIhAlaIhPheIhLeuIhLeuProIleIhLeuSerTyrIhMetIhAsnGlu	339
QY	1060	GGAATTAATCTGCAGTCTGTTTAAAGAGTTGAGGAACACAGTGTATTAATGTGACACA	1119
DB	340	GlyIhIleIhIhAlaSerGlnIhTyIhLeuGlnIhLeuGlnIhProGluAlaIleIleIleAlaPro	359
QY	1120	ACTCGAGAAATGTGCAACGATTTATTTGGAAGCCAGAAATTTCTTTGGACTTGT	1179
DB	360	ThryIhArgIhLeuIhIleAhnGlnIleIhTyIhLeuAhpIhArgIhysPheSerTyrGlyIhThCyS	379
QY	1180	GTAAGAGCTGTTGTTATATATGSGGGAACCCAGCTGGGACATTCATTCGACAAATAGTA	1239
DB	380	ValIhArgProValIhValIhTyIhGlyIhIleGlnIhProValIhIhAlaMetArgAhpValGlu	399
QY	1240	CAAGGCTGTATATATATATGTGCTACCTCGGAAGATGATGATATCATAGCAAGAA	1299
DB	400	LysGlyIhCysAhnIleIhLeuCyAlaIhThProGlyIhArgIhLeuIhSerIleValSerLysGlu	419
QY	1300	AAGATTGGTCTCAACAGATCAATTAAGTTGTTGATGAAAGCTGATCGCATTTGCAT	1359
DB	420	LysIleGlyIhLeuSerIhysLysAhnGlyIhLeuValIhLeuAhpGluIhAlaAhpArgMetIhAsp	439
QY	1360	ATGGGCTTTGGTCCAGAAATGAAGACTTAATTTCTTGCCCAAGATGCCATCAAGAA	1419

Db 440 MetGlyPheAlaProGluIleGluIleuMetThrIysProGlyMetProThrIleGlu 459

Qy 1420 CAGCGCCAAACCCCTTATGTCAGTGCAACTTTCCAGAGAAATCAAGGTTGGCTGCA 1479

Db 460 LysArgGlnThrLeuMetPheSerAlaThrIyrProGluGluIleArgGluAlaSer 479

Qy 1480 GAGTTTTTAAAGTCAAAATATATCTGTTTGTCTGTGGACAAGGGGTGGAGCATGAGA 1539

Db 480 AsnIyrLeuIysSerGlnIleLeuPheValIValGlyLeuValIleGlyAlaCysSer 499

Qy 1540 GATGTTTCAGCAGACCCGTCCTCCAAAGTTGGCCAGTCTCCAAAGAGAAAGAACTGTTGAA 1599

Db 500 AspValIleAlaGlnThrValIleuGlyMetArgGluMetGlyLysMetGlyLysLeuGlu 519

Qy 1600 ATTTCGCAAAACATAGGGGATGAAAGAACTATGCTTTTGTGTAACCTAGAAAAAGCA 1659

Db 520 IleLeuIysSerSerGlnIleGluArgThrMetIlePheValAlaThrIleLysIleAla 539

Qy 1660 GATTTTACTGCAACTTTTCTTGTCAAGAAAAATATGACTACAGATTCATGGTGAT 1719

Db 540 AspPheIleAlaGlyIyrLeuIleCysGlnIleLysPheSerSerThrSerIleAlaGlyAsp 559

Qy 1720 CGGGAACAGAGAGACGGGAGAGCAAGCTCTTGAGAGATTTGCTTTGGAAATGGCCAGTT 1779

Db 560 ArgGluGlnIyrGlnArgIleSerAlaLeuThrPhePheAlaGlnIyrGlyLysCysThrVal 579

Qy 1780 CTTTGTGCTACTCAGTAGCTCCAGAGGGCTGATATGTAATGTAATGTCACATGTTATC 1839

Db 580 IleValCysThrAlaValAlaIleAlaArgIleuAspIleGluMetValGlnIleValIle 599

Qy 1840 AATTGATATCTTCCTCTTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1899

Db 600 AsnIyrAspValIProIyGluValAspGlyIyrValIleAspGlyIyrGlnIleGlyIyrGly 619

Qy 1900 TGTGGGATFACCGGAGAGCAATTCCTTTTGTGCTGAATGGATTAACCATTTGCA 1959

Db 620 CysGlyAsnIleGlyLysAlaThrSerPhePheAsnValGlnAspAsnIleAlaIleAla 639

Qy 1960 CAGCCTTAGTAAAGATATGACAGATGCTCAACAGATGTTCTGTCAGTGTGGAAGAA 2019

Db 640 ArgProLeuValIleIleLeuThrAspAlaIleGlnIleValIProIatIyrPheGluGlu 659

Qy 2020 ATTGCTTTAGTACATTAATTCCTGCTTCAAGTGTAGTACAGAGCAAGCTGTTTGA 2079

Db 660 IleAla-----PheGlyGlnIleGlyAlaLeuAsnSerPheIyr 672

Qy 2080 TCAGTGTATAC-----AGAAAGGCAAGGCACTTTGAACACAGCTGGGTTTCT 2130

Db 673 AlaAlaAspSerMetGlyGluGlnIleAlaGlyLysAlaValIThrThrProSerPheAla 692

Qy 2131 TCTTCACGAGCTCCCAATCCAGTAGATGATGATGATGATGATGATGATGATGATGAT 2172

Db 693 GlnGlu-----GluGlnIleAspIyrPhe 700

RESULT 5

AS8768

ATP-dependent RNA helicase homolog - fruit fly (*Drosophila melanogaster*)

C/Species: *Drosophila melanogaster*

C/Date: 31-Dec-1990 #sequence revision 17-Apr-1998 #text\_change 19-Jan-2001

C/Accession: AS8768, S01676, S10129, A11922

R/Label: L.; Diehl-Jones, W.; Lasko, P.

Unpublished results, 1995, cited by GenBank in release 103.0

A/Description: Localization of Vasa protein to the *Drosophila* pole plasm is independent

A/Reference number: AS8768

A/Accession: AS8768

A/Status: not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 1-661 <LA81>

A/Cross-References: EMBL:Y12945; NID:G433675; PIDN:CAA1405.1; PID:G1054723

R/Lasko, P.F.; Ashburner, M.

Nature 335, 611-617, 1988

A/Title: The product of the *Drosophila* gene vasa is very similar to eukaryotic initiation

```

A:Reference number: S01676; MUID:89014721; PMID:3140040
A:Accession: S01676
A:Molecule type: DNA
A:Residues: 1-34, 'R', 36-448, 'LRF', 453-460, 'R', 462-589, 'Q', 591-661 <IAS2>
A:Cross-references: EMBL:X12945; NID:g433675
R:Ashburner, M.
submitted to the EMBL Data Library, October 1988
A:Reference number: S10129
A:Accession: S10129
A:Molecule type: DNA
A:Residues: 1-34, 'R', 36-644, 'Q', 645-661 <ASH>
A:Cross-references: EMBL:X12945; NID:g433675
R:Hay, B.; Jan, L.Y.; Jan, Y.N.
Cell 55, 577-587, 1988
A:Title: A protein component of Drosophila polar granules is encoded by vasa and has exte
A:Reference number: A31922; MUID:89028669; PMID:3052853
A:Accession: A31922
A:Molecule type: RNA
A:Residues: 1-34, 'R', 36-154, 168-264, 'Y', 266-321, 'C', 323-451, 'F', 453-581, 'R', 583-593, 'H', 'E', 'E'
A:Cross-references: GB:M2560; NID:g158795; PIDN:AAA29013.1; PID:g158796
A>Note: The authors translated the codon TGI for residue 322 as Val
C:Genetics:
A:Gene: vasa
A:Cross-references: FlyBase:FBgn0003970
A:Introns: 8/3; 138/1; 234/3; 482/3; 554/1; 644/2
C:Superfamily: ATP-dependent RNA helicase DBP1
C:Keywords: ATP; nucleotide binding; P-loop
F:289-296/Region: nucleotide-binding motif A (P-loop)
F:395-400/Region: nucleotide-binding motif B
F:399-402/Region: DEAD motif

Alignment Scores:
Pred. No.: 2.15e-83 Length: 661
Score: 1274.00 Matches: 314
Percent Similarity: 56.73% Conservative: 103
Best Local Similarity: 42.72% Mismatches: 228
Query Match: 33.22% Indels: 90
DB: 2 Gaps: 25

US-09-714-865-15 (1-2172) x A58768 (1-661)
QY 10 GAAGATTGGGAGCAGCAATCAACCTCATATGTTCTTATGTTCCCATATTGAG-- 66
Db 3 AspAspLTPAspAspGlu-----ProIleValAspThr 13
QY 67 AAGCATAGGTATTTCTGGAGAAATGAGACCAATTTTAAACAGACTCCAGCTTCATCA 126
Db 14 ArgGlyAlaArgGlyGlyIleAspTrpSerAspAspGluAspThrAlaIleYsserPheSecGly 33
QY 127 GAATATGAT--GATGGACCTTCTCGAAGAGATCATTTCTGAAAGTGGATTGGCTCT 183
Db 34 GlnAlaGluGlyAspGlyAlaGlyGly-----SerGlyGlyGluGly 48
QY 184 GGGCGGAATTTTGAAGACAGAGATCTGAGTGTATTAAGCAGATATATACACCA 243
Db 49 GlyTyrgInGlyGlyAsnTrpAsp-----Val 57
QY 244 ATGGGTGTTTTGAGATTGGAAGAGTTTGGAAACAGAGTTTTCAAACAGAGCTTT 303
Db 58 PheGlyArgGlyIleGlyGlyArgGlyGlyAlaGlyGlyIleTyraGlyGlyAsnArg 77
QY 304 GAAGATGCTGATAGCTCTGCTTTCTGGAGAGATCTAAGTATGATCACTCCGAAGATATCA 363
Db 78 AspGlyGlyGlyPheHisGlyGlyArgArgGlyGlyGluIleAspPhe----- 93
QY 364 ACACGGAACAGAGGGGTTTCCAAGAGAGCGGCTATCGAATGAAATTAATTCAAGACT 423
Db 94 -----ArgGly-----GlyGluGlyGlyPheArgGlyGlyGlnGlySerArg 108
QY 424 TCAGGGCCA---TACAGAGAGGTGGAAGAGTATGTTCCGAGTTCCGTGAGAGATTT 480
Db 109 GlyGlyGlnGlyGlySerArgGlyGlyGlnGlyGlyPheArgGlyGlyGlnGlyGlyPhe 128

```

QY 481 ---GCTCTAGAGAGTCCAAATTAATGACTTGACCCAGACGAATGATGACGCCACTGCT 537  
 DB 129 ArgGlyArgLeuTyrGluAsnGluAspGlyAsp-----GluArgArgGly 143  
 QY 538 GGCCTT-----TTTGGTTCTAGAAAGACCATTAATTAAGTGCACACAGTAAT 582  
 DB 144 ArgLeuAspArgGluGluArgGlyGluArgGlyValArgLeuAsp----- 159  
 QY 583 GGTGATACCTCTCAAGACAGAGTGCAGTGAAGTGAACGA-----GGTGGTAC--- 633  
 DB 160 -----ArgGluGluArgGlyGluArgGlyGluArgGlyValArgGlyPheAla 176  
 QY 634 ---AAAGCTTAATGAAGAATPAATACAGGCTCTGGAAGAATCTTGGAAGTCAAA 690  
 DB 177 ArgArgArgGlnGluAspAspIle-----AsnAsnAsnAsnAsnIle 191  
 QY 691 GCAGAAGAGGAAGAAAGTACTGATCTCAAGACCAAAAGTACATCAATCCCTCCT 750  
 DB 192 AlaGlu-----AspValGluArgGlyArgGluPheTyrIleProProGlu 206  
 QY 751 CCACCTGAGATGAGAGACTCCATCTTTCACAT---TATCAGACAGGCAATTAATTCGAC 807  
 DB 207 ProSerAsnAspAlaIleGluIlePheSerSerGlyIleAlaSerGlyIleHisPheSer 226  
 QY 808 AAATGACACATATCTTGTGAAGTGTCTGACATGATGACACACCAAGATTTGACT 867  
 DB 227 LysTyrAsnAsnIleProValIleValIleThrGlySerAspValProGlnProIleGlnHis 246  
 QY 868 TTGTAAGAAAGCTATCTCTGACACCTGAATACAAACATTTGCTAAGCTGTTTACT 927  
 DB 247 PheThrSerIleAspLeuAspIleIleIleAspAsnValAsnIleSerGlyPheLys 266  
 QY 928 AACCTACTCTCTGCAAAATACAGATTCCTATCTATCTATCTAGACAGACAGATTTGATG 987  
 DB 267 IleProThrProIleGlnLysCysSerIleProValIleSerSerGlyArgAspLeuMet 286  
 QY 988 GCTTGTGCTCAAAACAGGCTGGAAGACTCGCGCTTTTCTCTCAATTTTGACTCAT 1047  
 DB 287 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuProIleLeuSerLys 306  
 QY 1048 ATATGATGATGATGATTAATCTGCCAGTCTGTTTAAAGATTGACAGACCAAGTGTATT 1107  
 DB 307 LeuLeuGluAsp-----ProHisGluLeuGluGluArgProGlnValVal 322  
 QY 1108 ATTGTGACCAACTGAGAAATGTTGTCACACCAATTAATTTAAGGACAGAAATTTCT 1167  
 DB 323 IleValSerProThrArgIleGluAlaIleGlnIlePheAsnGluAlaArgLysPheAla 342  
 QY 1168 TTGTGACTGTGTAGAGCTGTGTATATATATGAGGGAACCCAGCTGGACATTCATTT 1227  
 DB 343 PheGluSerTyrLeuLysIleGlyIleValIleValIleThrIleSerPheArgHisGlnAsn 362  
 QY 1228 CGCAAAATGATACAGGCTGTAATATATATATGTCCTCTGAGAGACTGATGATATC 1287  
 DB 363 GluCysIleThrArgGlyCysHisValValIleAlaThrProGlyArgLeuLeuAspPhe 382  
 QY 1288 ATAGGCAAAAGAAAGATGTTGTCCAACAGATCAAAATCTAGTTTGGATGAAGCTGAT 1347  
 DB 383 ValAspArgThrPheIleThrPheGluAspThrArgPheValIleLeuAspGluAlaAsp 402  
 QY 1348 CGCATGTGATATGCTTTGCTCCAGAAATGAAGAATTAATTTCTCCAGAGATG 1407  
 DB 403 ArgMetLeuAspMetGlyPheSerGluAspMetArgArgIleMetThrHisValThrMet 422  
 QY 1408 CCATCAAGAACACGCCCAACCTTATGTTGTCGCAACTTTCCAGAGAAATTCAA 1467  
 DB 423 -----ArgProGlnHisGlnIleThrLeuMetPheSerIleThrPheProGluGluIleGln 440  
 QY 1468 AGGTGGCTCAGAGCTTTTAAAGTCAATTAATCTGTTGTGCTGAGCAAGTGGCT 1527  
 DB 441 ArgMetAlaIleGluPheLys---AsnTyrValSerValAlaIleGlyIleValGly 459  
 QY 1528 GGAGCATGTAGAGATGTTGACAGACCGTTCTCCAAAGTTGCGCAATTCAGAAAGAGAA 1587

DB 460 GlyAlaCysSerAspValIleGlnThrIleTyrGluValAsnIleTyrAlaLysSer 479  
 QY 1588 AACCTGTTGAATTTCTGCGAAACATAGGGGATGAAGAAGACTATGCTTTGTTGAACT 1647  
 DB 480 LysLeuIleGluIleLeuSerGluGlnAlaAsp---GlyThrIleValPheValGluThr 498  
 QY 1648 AAGAAAAAGACAGATTTTACTGCACTTTCTTGTCTCAAGAAAAATATCACTACAAGT 1707  
 DB 499 LysArgGlyAlaAspPheLeuAlaSerPheLeuSerGluLysGluPheProThrIleSer 518  
 QY 1708 ATCCATGGTGATGGGAACACAGACAGCGGAGCAAGCTCTTGAGATTTTCCTTTGA 1767  
 DB 519 IleHisGlyAspArgLeuGlnSerGlnArgGluGlnAlaLeuArgAspPheLysAsnGly 538  
 QY 1768 AAGTGCACAGTCTTGTGCTGATCTAGTACAGTCCAGAGCGGCTGATATGAATGTC 1827  
 DB 539 SerMetLysValLeuIleAlaThrSerValAlaSerArgGlyLeuAspIleLysAsnIle 558  
 QY 1828 CAACATGTTATCAATTTTGTATCTTCTTACCATTTGATGAATGATTCGAATTTGG 1887  
 DB 559 LysHisValIleAsnTyrAspMetProSerLysIleAspAspTyrValHisArgIleGly 578  
 QY 1888 CTTACTGGTCTGTTGGGAATACTGSCAGACAAATTTCTTTTGTGATTCGAT 1947  
 DB 579 ArgThrGlyCysValGlyAsnAsnGlyArgGlaThrSerPhePheAspProGluLysAsp 598  
 QY 1948 AACCATTTGACACAGCTCTAGTAAAGTATTCACAGATGCTCAACAGCATGTTCCGCA 2007  
 DB 599 ArgAlaIleAlaAlaAspLeuValLysIleLeuGluGlySerGlyGlnThrValProAsp 618  
 QY 2008 TCGTTGGAAGAAATTCCTTTAGTACATCATCTCTGCTTCACTGTTAGTAACAGAGA 2067  
 DB 619 PheLeuArg-----ThrCysGlyAlaGlyValAspGlyTyrSerAsn 633  
 QY 2068 AACGTGTTGCATCAGTTGATACCAAGAGGCAAGACCTTTGAACACAGCTGGCTT 2127  
 DB 634 GlnAspPheGlyGlyValAspValArg---GlyArgGly-----AsnTyrValGly--- 649  
 QY 2128 TCTTCTTACAGAGCTCCCAATCCAGTATGATGATGATCAGGAT 2172  
 DB 650 -----AspAlaThrAsnValGluGluGluGluGlnIleThrAsp 661  
 RESULT 6  
 184741  
 RNA helicase - mouse  
 N/Alternate names: RNA helicase ERH  
 C/Species: Mus musculus (house mouse)  
 C/Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 19-Jan-2001  
 C/Accession: 184741; S56112  
 R/Gee, S.L.; Comboy, J.G.  
 Gene 140, 171-177, 1994  
 A/Title: Mouse erythroid cells express multiple putative RNA helicase genes exhibiting hu  
 A/Reference number: 149731; PMID:94192995; PMID:8144024  
 A/Accession: 184741  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-662 <RES>  
 A/Cross-references: GB:125126; NID:9407995; PIDN:AAAS3630.1; PID:9407996  
 A/Genetics: RES1  
 A/Sowden, J.; Putt, W.; Morrison, K.; Beddington, R.; Edwards, Y.  
 Biochem. J. 308, 839-846, 1995  
 A/Title: The embryonic RNA helicase gene (ERH): a new member of the DEAD box family of R  
 A/Reference number: S56112; PMID:97104282; PMID:8948440  
 A/Accession: S56112  
 A/Molecule type: mRNA  
 A/Residues: 1-229, 'R', 231-662 <SOM>  
 A/References: GB:238117; NID:91835121; PID:91835122  
 A/Experimental source: strain c57bl/6; notochord  
 A/Genetics: SOM1  
 A/Note: the sequence is revised in GenBank entry M26BRNHL, release 117, (PID:1835122)  
 A/Note: the revised sequence is now identical to PIR accession 184741  
 C/Genetics: <RES1>











```

|||||
Db 343 LeuValLeuAspGluAlaAspArgMetLeuAspMet1yPheGluProGlnIleArgArg 362
QY 1387 TTAAATTTCTCCAGGATGCCATCAAGGAACAGCGCAACCCCTTATTCAGTGCA 1446
Db 363 IleValGluGlnAspThrMetProProlyGlyAlaArgHisThrMetMetPheSerAla 382
QY 1447 ACTTTTCAGAGAAATTCAGAGGTTGGCTGCAGAGTTTAAAGTCAAAATATCTGTT 1506
Db 383 ThrPheProlyGlnIleGlnMetLeuAlaArgAspPheLeu---AspGluTyrIlePhe 401
QY 1507 GTTGCTGTTCGACAAAGTGGTGAGCATGTAGAGATTTCAACGACCGCTTCCAAATT 1566
Db 402 LeuAlaValGlyArgValGlySerThrSerGlnAsnIleThrGlnIleValIleTyrVal 421
QY 1567 GGCACGTTCTCAAAAAGAAAGCTCGTGAATTCGTGGAACATATAGG---GATGAA 1633
Db 422 GluGluAlaAspIleAspSerPheLeuLeuAspLeuAsnAlaThrGlyLysAspSer 441
QY 1624 AGAAGTATGGTCTTTTGTGAACATAGAAAAAGCAGATTTTACTGCACCTTTCTTGT 1683
Db 442 LeuIleLeuValIlePheValGluThrIleLysGlyAlaAspSerLeuGluAspPheLeuTyr 461
QY 1684 CAAGAAAAATATCACTAACAATGATATCATGTGATGGGAAACAGAGAGCGGAGCAA 1743
Db 462 HisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGlu 481
QY 1744 GCTCTTGAGATTTGGCTTTGGAAGGCCAGCTTGTGTTGCTACTTACAGTAGCGCC 1803
Db 482 AlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAla 501
QY 1804 AGAGGGCTGATATTGAAAAATGTGCAACATGTTATCAATTTGATCTTCTTACCAT 1863
Db 502 ArgGlyLeuAspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIle 521
QY 1864 GATGATATGTTATCGAATTTGGCGTACTGGTGTGTTGGAAATCTGCGACAGCAATT 1923
Db 522 GluGluTyrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThr 541
QY 1924 TCCCTTTTATCTGTGATGGAATGAATACCATTTAGCACAGCCCTCAGTAAAGTATTGACA 1983
Db 542 SerPhePheAsn---GluArgAsnIleAsnIleThrLysAspLeuLeuVal 560
QY 1984 GATGCTCAACAGATGTTCTCGATGTTGGAAGAAATTCCTTATGACATACATTCT 2043
Db 561 GluAlaLysGlnGluValProSerTyrLeuGluAsnMetAlaPheGlnHis----- 578
QY 2044 GGCTTCACTGATGATACAGAGAAACGTG-----TTTGATCA 2082
Db 579 ---TyrIleGlySerArgGlyArgSerLysSerArgPheSerGlyGlyPheGlyAla 597
QY 2083 GTTGATCCAGAAAGGCGCAAGACACTTTGAACACACTGGTTTCTTTCACGAGCT 2142
Db 598 ArgAspIleArgGln---SerSerGlyAlaSerSerSerPheSerSerGlyArgAla 616
QY 2143 CCAAT 2148
Db 617 SerAsn 618

RESULT 9
T48796
Probable ATP-dependent RNA helicase DED1 (imported) - Neurospora crassa
N;Alternate names: protein 15B6.40
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 17-Nov-2000
R;Schulte, U., Aign, V., Hohelsiel, J., Brandt, P., Fartmann, B., Holland, R., Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Accession: T48796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-688 <SCH>

```

```

A;Cross-references: EMBL:ALJ53822; GSPDB:GN00112; NCSP:15B6.40
A;Experimental source: cosmid contig 15B6; strain 74
C;Genetics:
A;Gene: NCSP:15B6.40
A;Map position: 2
A;Intons: 63/2
C;Superfamily: ATP-dependent RNA helicase DBP1

Alignment Scores:
Pred. No. 2,33e-76 Length: 688
Score: 1176.00 Matches: 265
Percent Similarity: 60.74% Conservative: 80
Best Local Similarity: 46.65% Mismatches: 177
Query Match: 29.74% Indels: 46
DB: Gaps: 11

US-09-714-865-15 (1-2172) x T48796 (1-688)
QY 409 AATTAATTCAGAACCTTCAGGCCATACAGAAAGGTGGAAGAGTATGTTCCAGAGTTGC 468
Db 65 AsnAsnTyrAspAlaArgGlyPro-----GlyGlyGlyLysAsnTyrProAlaPro 81
QY 469 CGTGAGAGATTGGCTTAGGAAGTCCAATATGACTTAGACCCAGACGATATGACAG 528
Db 82 GlyGlyProProGlyProGlyPheGlyGly----- 91
QY 529 CGCACTGGTGGCCCTTTTGGTTTCTAGAACGACGATTTAAGTGCACAGTAATGCTGAT 588
Db 92 GlnGlnGlyAlaGlyTyrGlyGlyProArgPro-----GlnGlyGlyPheAsnProAsn 109
QY 589 ACTTCTCAAGCAGAAAGTGGCAGTGAAGTGAACGAGGTGTTAACAAGGTTTAAATGAA 648
Db 110 AlaTyrArgGlyAsnAlaGlyAlaGlyAlaGlyAlaGlyAlaGlyGlyGlySer 129
QY 649 GAAGTAATTAACAGCCTCTGGAAGAAATTTCTGGAAGTCAAGACGACAGAGAGAAAGT 708
Db 130 PheSerAsnArgIleSerGlyAspGlyGlnTyrArg-----AspGly----- 143
QY 709 AGTGATATCTCAAGACCAAAAGTACCTCATGCC---CTTCCTCAGCTGAGATGAG 765
Db 144 -----LysHisIleProGlyProAlaAsnProArgValGlu 155
QY 766 GACTCATCTTT-----GCACATTATCAGACAGGCAATAACTTCGAC 807
Db 156 ArgGluLeuPheGlyThrProAspAspProSerLysGlnHisThrGlyIleAsnPheGlu 175
QY 808 AAATAGACACTATTCTGTGGAAGTGTGACATGATGACACACGACCAATTCGACT 867
Db 176 LysTyrAspAspIleProValGluAlaSerGlyAspAsnValProGluProValLeuThr 195
QY 868 TTGGAAGAGCTAATCTCTGTGACACCTGAATTAACAACATTCCTAAGCTGTTTACT 927
Db 196 PheSerAsnProProLeuAspAsnHisLeuIleSerAsnIleGlnLeuAlaArgTyrAsn 215
QY 928 AAGCTTACTCTCTGCAAAAATACAGTATTCATATATCTTACATGACAGAGATTTGATG 987
Db 216 ValProThrProValGlnTyrSerIleProIleValMetGlyAlaArgAspLeuMet 235
QY 988 GCTTGCTCAACAGAGTCTGGAAAGACTCGGCGCTTTCTCTTACCAATTTGGCTCAT 1047
Db 236 AlaCysAlaGlnThrGlySerGlyLysThrGlyGlyPheLeuPheProIleLeuSerGln 255
QY 1048 ATGATGATCATGATGA-----ATPACTGCCAGTCGTTT-----AAA 1083
Db 256 SerPheHisThrIleProSerProIleProAlaSerAlaAlaGlyAlaTyrGlyArgGln 275
QY 1084 GAGTTGCAGGAACCAAGATGATATTATTTAGACCAACATCGAAGATTTGCAACAGATT 1143
Db 276 ArgLysAlaTyrProThrAlaLeuIleLeuAlaProThrArgGluLeuValSerGlnIle 235
QY 1144 TATTGGAAGCCAGAAATTTCTTTGGGACTTGTGAAGAGCTGTTGTTATATATGGG 1203
Db 296 TyrAspGluAlaArgLysPheAlaTyrArgSerTyrValArgProCysValValTyrGly 315

```

```

QY 1204 GGAACCCAGCTGGACATTCATTCGAAATAGTACAGGCTGTAATATATATGTCCT 1263
DB 316 G1YAlaAspIleGlySerGlnLeuArgInIleGluArgGlyCysAspLeuLeuValAla 335
QY 1264 ACTCTCGAAGACATGATGATATATATGCGAAGAAAGAAAGATGCTCTCAACAGATCAA 1323
DB 336 ThrProGlyArgLeuValAspLeuIleGluArgGlyArgIleSerLeuCysAsnIleLys 355
QY 1324 TACTTACTGTTGATGAAAGCTGATCGATGTTGGATATGCTGTTGGTCCGAAATGAG 1383
DB 356 TyrLeuValLeuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArg 375
QY 1384 AAGTAAATTTCTTCCCGACGAAATGCTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1443
DB 376 ArgIleValIleGluGlyGluAspMetProLysValAsnAspArgGlnIleLeuMetPheSer 395
QY 1444 GCAACTTTTCCAGAGAAATTCAGAAAGTTGCTGCGACAGTTTAAAGTCAAAATATCTG 1503
DB 396 AlaThrPheProArgAspIleGlnIleLeuAlaArgAspPheLeuLys--AspTyrIle 414
QY 1504 TTGTTGCTGTTGGAAGTGGGTGAGCATGTAGAGATGTTGACGACGCTTCTCAA 1563
DB 415 PheLeuSerValIleArgValIleGlySerThrSerGluSerIleThrGlnLysValIleTyr 434
QY 1564 GTTGCCAGCTCTCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1623
DB 435 ValGluAspIleAspLysArgSerValLeuLeuAspIleLeuHisThrHisAlaGlyGly 454
QY 1624 AGAAGTATGCTCTTGTGTAAGTAAAGAAAGAAAGAAAGAAAGTCTTCTGCAATTTCTTGT 1683
DB 455 LeuThrLeuIlePheValIleGluThrLysArgMetAlaAspSerLeuSerAspPheLeuIle 474
QY 1684 CAAGAAAAAATATCACTACAGATTCATGCTGATGCTGGAACAGAGAGACGGGAGCAA 1743
DB 475 AsnGlnAsnPheProAlaThrSerIleHisGlyAspArgThrGlnArgGluArgGluArg 494
QY 1744 GCTCTGGAGATTTTTCGCTTGGAAAGTCCGACGCTTCTGCTGCTTCTGCTGCTGCTGCC 1803
DB 495 AlaLeuGluMetPheArgAsnGlyArgCysProIleLeuValAlaAlaValAlaAlaAla 514
QY 1804 AGAGGCTGATATATGAAATATGTCACATGATTCATATTCATTTCTTCTTCAACAT 1863
DB 515 ArgGlyLeuAspIlePheProAsnValThrHisValIleAsnLysArgAspLeuProThrAspIle 534
QY 1864 GATGATATGTTTCATCGAATTTGGCGTACTGCTGCTTGGGAATCTGCGAGAGCAATT 1923
DB 535 AspAspTyrValHisArgIleGlyArgThrGlyArgAlaGlyAsnThrGlyIleAlaThr 554
QY 1924 TCCTTTTTCATCTGGAATCGAATACATTTAGACAGCTGCTTGAATTAAGTATTGACA 1983
DB 555 AlaPhePheAsn--ArgGlyAsnArgGlyValValAlaArgGluLeuLeuGluLeuLys 573
QY 1984 GATGCTCAACAGATGTTCTCGATGTTGGAAGAAATGCTTGAATACATATTCCT 2043
DB 574 GluAlaAsnGlnGluValProAlaPheLeuGluThrIleAlaArgGluSerSerPheGly 593
QY 2044 GCGTTCACTGCTAGTACAGAGCA 2067
DB 594 GlyGlyArgGlyGlyArgGlyGly 601

```

## RESULT 10

S13653  
 ATP-dependent RNA helicase DED1 - yeast (*Saccharomyces cerevisiae*)  
 N/Alternate names: protein O4836; protein YOR204w  
 C/Species: *Saccharomyces cerevisiae*  
 C/Date: 16-Sep-1992 #sequence revision 16-Sep-1992 #text\_change 19-Jan-2001  
 C/Accession: S13653; S07683; S67096  
 R/Jamieson, D.J.; Rahe, B.; Pringle, J.; Beggs, J.D.  
 Nature 349, 715-717, 1991  
 A/Title: A suppressor of a yeast splicing mutation (prp8-1) encodes a putative ATP-depen  
 A/Reference number: S13653; MUID:91141585; PMID:1996139  
 A/Accession: S13653

A/Molecule type: DNA  
 A/Residues: 1-604 <JAM>  
 A/Cross-references: EMBL:X57278; NID:93646; PIDN:CAA40546.1; PID:93647  
 A/Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 37-Met &  
 R/Struhl, K.  
 Nucleic Acids Res. 13, 8587-8601, 1985  
 A/Title: Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* ge  
 A/Reference number: S07681; MUID:86093663; PMID:3001645  
 A/Accession: S07683  
 A/Molecule type: DNA  
 A/Residues: 1-36, 'M', 38-113 <STR>  
 A/Cross-references: EMBL:X03245; NID:93778; PIDN:CAA27004.1; PID:93781  
 R/Hughes, B.; Pohl, T.M.  
 Submitted to the Protein Sequence Database, July 1996  
 A/Reference number: S66685  
 A/Accession: S67096  
 A/Molecule type: DNA  
 A/Residues: 1-604 <HUG>  
 A/Cross-references: EMBL:275110; NID:91420476; PIDN:CAA99419.1; PID:91420479; MIPS:YOR204  
 A/Experimental source: strain S288C  
 C/Genetics:  
 A/Gene: SCD; DED1; SPP81  
 A/Cross-references: SGD:S0005730, MIPS:YOR204w  
 A/Map position: 15R  
 C/Superfamily: ATP-dependent RNA helicase DBP1  
 C/Keywords: ATP; nucleotide binding; P-loop  
 F:186-193/Region: nucleotide-binding motif A (P-loop)  
 F:302-307/Region: nucleotide-binding motif B  
 F:306-309/Region: DEAD motif

Alignment Scores:  
 Pred. No.: 8.02e-73 Length: 604  
 Score: 1126.50 Matches: 261  
 Percent Similarity: 56.66% Conservative: 92  
 Best Local Similarity: 41.89% Mismatches: 187  
 Query Match: 28.49% Indels: 83  
 Gaps: 16

US-09-714-865-15 (1-2172) x S13653 (1-604)

```

QY 337 TCTAGTATGATCTGCGAAGTAAAT-----CCAAACAGCAAGAGGCTTTCCAG 387
DB 12 SerIleAsnAspAsnGlnAsnGlyTyrValProProHisLeuArgGlyLysProArg 31
QY 388 AGAGCGCGCTATGAGATGGAATTAATTCAGAAAGCTTCAAGGCGCATTCAGAGAGT--- 444
DB 32 SerAlaArgAsnAsnSerSerAsnTyrAsnAsnAsnGlyLysValAsnGlyLysArg 51
QY 445 GGAAGAGTAGTTTC-----CGAGTTGCGGTGAGAGATTTGGTCTAGAGAGTCCAAAT 498
DB 52 GlyGlyGlySerPhePheSerAsnAsnArgArgGlyLysTyrGly----- 66
QY 499 AATGACTTAAACCCAGCAGATGATGTCAGCGCACTGGTGGCTTTTGGTCTTAGAGA 558
DB 67 -----AsnLysIlePhePheGly----- 72
QY 559 CCAATTAAAGTGCACAGGTAATGCGATACCTTCGAAAGCAAGAGTGCAGTGAAGT 618
DB 73 -----GlyAsnAsnGlyGlySerArgSerAsnLysArgSerGly----- 85
QY 619 GAACGAGGTGTTACAAAGGTTTAATGAAGAAATTAACAGGCTTCGAAAGAAATTC 678
DB 86 -----GlyArg 87
QY 679 TCGAAGTCAAGACAGAGAGAGAGAAAGTAGTATCTCAAGGACCAAAAGTACCTAC 738
DB 88 Trp-----IleAspGly-----LysHis 93
QY 739 ATACCCCTCTCCACCTGAGAGT---GAGGACTCATCTTTGCA----- 780
DB 94 ValProAlaProArgAsnGluLysAlaGluIleAlaIlePheGlyValProGluAspPro 113
QY 781 CATATACG---ACAGCATTAACCTTCGACAAATACGACATATATCTTGTGAGAGTCT 837

```

```

Db      114 AsnphelInserSerGlyIleasnPhaSpasnIyrAspIleProValAlaSer 133
      838 GGACATGATGACACACAGCAATTCGACTTGAAGAAGTAATCTGTGCACACTG 897
      134 G|LysaspValProGluProIleTherGluPherSerProProLeuaspGlyLeu 153
      898 AATTAACAACATGCTTAAGCTGTTATAGTAAGCTTACTCTGTGCAAAATACATATT 957
      154 LeuGluasnIleLysLeuAlaArgPheThrLysProThrProValGlnLysSerVal 173
      958 CCGATCATCTTGACAGACAGATTTGATGCTGTCTGCAAAAGAGCTCTGGGAAGCT 1017
      174 ProIleValAlaasnGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThr 193
      1018 GCGGCTTTCTCCCTCAATTTGCTCATATGATGATGATGATGATGATGATGATGATGAT 1068
      194 G|GlyLysPheLeuPheProValLeuSerGluSerPheLysThrGlyProSerProGluPro 213
      1069 -----GCCAGTCCTTTTAAAGCTGACAGAACCCAGAGTGTATATTTATGACACCA 1119
      214 GluSerGlnGlySerPheThrGlnArgLysAlaLysProThrAlaValIleMetAlaPro 233
      1120 ACTCGAATTTGGTCAACAGATTTATTTGGAAGCCAGAAATTTCTTTGGACCTGT 1179
      234 ThrArgGluLeuAlaThrGlnIlePheAspGluAlaLysLysPheThrLysArgSerTrp 253
      1180 GTAGAGCTGTTGTTATATATGAGGGGAGACCCAGCTGGACATTCATTCAGCAAAATGTA 1229
      254 ValLysAlaCysValValIyrGlyGlySerProIleLysGlnLeuAlaValGlnIleGlu 273
      1240 CAAGGCTGTAATATATATGCTACTCTCGAAGACTGATGATCATCAGCAAAAGAA 1299
      274 ArgGlyLysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuGluArgGly 293
      1300 AAGATTGCTGCAACAGATCAAACTTACTGTTGATGATGATGATGATGATGATGATGAT 1359
      294 LysIleSerLeuAlaasnValIyrLysLysValLeuAspGluAlaAspArgMetLeuAsp 313
      1360 ATGGGTTTGGTCCAGAAATGAGAAAGTAATTTCTGCCAGAAATGCCATCAAGAA 1419
      314 MetGlyPheGluProGlnIleArgHisIleValAlaLysCysAspMetThrProValGly 333
      1420 CAGCGCCAAACCTTATGTCAGTGAACCTTTTCAGAGAAATTCAGAGGTTGGTGA 1479
      334 GluArgGlnThrLeuMetPheSerAlaThrPheProAlaAspIleGlnIleValAlaArg 353
      1480 GAGTTTAAAGTCAAAATTAATCTGTTGCTGCTGTTGACAGAGTGGTGAAGCATGTAGA 1539
      354 AspPheLeu--SerAspLysIlePheLeuSerValGlyArgValGlySerThrSerGlu 372
      1540 GATGTTGACAGACCCGTTCCAAAGTGGCCAGTTCCAAAAGAGAAAGCTGTTGAA 1599
      373 AsnIleThrGlnLysValLeuLysValGluAsnGlnAspLysLysSerAlaLeuLeuAsp 392
      1600 ATTCTCGCAAAACATAGAGGATGAAAGAACTATGCTTTGTTGAACATAAGAAAGCA 1659
      393 LeuLeuSerIleAspThrAspGlyLeuThrLeuIlePheValGlnThrLysArgMetAla 412
      1660 GATTTTACTGCAACTTTCTTTGTCAGAAAGAAATATCAATCAAGATTCATGATGAT 1719
      413 AspGlnLeuThrAspPheLeuIleMetGlnAsnPheArgAlaThrAlaIleHisGlyAsp 432
      1720 CGGGAACAGAGAGCGGAGCAAGCTCTTGAGATTTTGGCTTTGGAAAGTCCGCA 1779
      433 ArgThrIleSerGluArgGluArgAlaLeuAlaAlaPheArgSerGlyAlaAlaThrLeu 452
      1780 CTGTGCTACTTCAAGTGCAGAGGCTGATATTTGAATGTAAGCAATGATATC 1839
      453 LeuValAlaThrAlaValAlaAlaArgGlyLeuAspIleProAsnValThrHisValIle 472
      1840 AATTTGATCTCTTACCATGATGAATATGTTCAATGAAATGAGGCTGATGCTGCT 1899

```

```

Db      473 AsnIyrAspLeuProSerAspValAspAspLysValHisArgIleGlyArg 492
      1900 TGTGGAAATCTGCGACAGCAATTTCTTTGATCTTGAATCGATAACCAATTTAGCA 1959
      493 AlaGlyAsnThrGlyLeuAlaThrAlaPhePheAsnSerLysAsnSerAlaIleVal 511
      1960 CAGCCTTCTAATAAGTATTGACAGATGCTCAACAGAGATTTCTGATGTTGGAAGAA 2019
      512 LysGlyLeuHisGlnIleLeuThrGlnAlaAsnGlnGlnValProSerPheLeuLysAsp 531
      2020 ATTGCTTTAGTACATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2079
      532 AlaMetSerAla-----ProGlySerArgSerAsnSerArgArgGlyGlyPheGly 549
      2080 TCAGTGTATCCAGAAAGGCGACAGACACT-----TTGAACACAGCTGGGTTTCTCT 2133
      550 ArgAsnAsnAsnAspAspLysArgLysAlaGlyValAlaSerAlaGlyLysPheGlySer 569
      2134 TCAGAGACT 2142
      570 SerArgSer 572

RESULT 11
63343
probable ATP-dependent RNA helicase [similarity] - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: suppressor of uncontrolled mitosis
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence revision 21-Jan-2000 #text_change 19-Jan-2001
C:Accession: T43543; T41132; T43658; T50462; T43347
R:Forbes, K.L.C.; Enoch, T.
submitted to the EMBL Data Library, March 1998
A:Description: A suppressor of fission yeast checkpoint mutants encodes a putative ATP-dep
A:Reference number: Z22562
A:Accession: T43543
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-636 <FOR>
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.
submitted to the EMBL Data Library, April 1998
A:Reference number: Z21971
A:Accession: T41132
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-636 <LYN>
A:Cross-references: EMBL:AL022598; PIDN:CA18646.1; GSPDB:GN00068; SPDB:SPCC1795.11
A:Experimental source: strain 972n(-); cosmid C1795
R:Gallert, B.; Kearsey, S.E.; Lemhard, M.; Carlson, C.R.; Nurse, P.; Boye, E.; Labib, K
submitted to the EMBL Data Library, March 1999
A:Description: A putative RNA helicase reveals links between translation and the fission
A:Reference number: Z22607
A:Accession: T43658
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-636 <GRA>
A:Cross-references: EMBL:AJ237697; PIDN:CA840192.1
R:Li, H.Y.; Walworth, N.C.
submitted to the EMBL Data Library, August 1998
A:Description: A multi-copy suppressor of cdc2-14, Dep1, is homologous to the Sacchromyces
A:Reference number: Z25074
A:Accession: T50462
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-636 <LIU>
A:Cross-references: EMBL:AF084222; PIDN:AAC34121.1
R:Kawamukai, M.
Biochim. Biophys. Acta 1446, 93-101, 1999
A:Title: Isolation of a novel gene, moc2, encoding a putative RNA helicase as a suppressor
A:Reference number: Z22361; PMID:9936140; PMID:10395822
A:Accession: T43347
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-40, 'R', 42-636 <RAW>

```



QY 2146 -----AATCCAGTAGATGATGATCAGT 2169  
 Db 623 HieglyAsnThrTyraSerGlySerAlaGlnSerIrrp 635

## RESULT 12

626003  
 Probable ATP-dependent RNA helicase DBP1 - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: probable RNA helicase CA1; protein LPH8C; protein YPL119C  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 10-Apr-1996 #sequence revision 19-Apr-1996 #text\_change 19-Jan-2001  
 C:Accession: S62003; S16790; A34848  
 R:Schlenstedt, G.; Silver, P.A.  
 Submitted to the EMBL Data Library, December 1995  
 A:Reference number: S61996  
 A:Accession: S62003  
 A:Molecule type: DNA  
 A:Residues: 1-617 <SCH>  
 A:Cross-references: EMBL:U35503; NID:G1163087; PIDN:AAB68243.1; PID:G1163095; MIPS:YPL119C  
 R:Jamieson, D.J.; Beggs, J.D.  
 Mol. Microbiol. 5, 805-812, 1991  
 A:Title: A suppressor of yeast sps81/ded1 mutations encodes a very similar putative ATP-  
 A:Reference number: S16790; MUID:91312117; PMID:1857205  
 A:Accession: S16790  
 A:Molecule type: DNA  
 A:Residues: 1-42, 'RS', 45-47, 'K', 49-87, 'R', 89-114, 'OK', 116-617 <JAM>  
 A:Cross-references: EMBL:X55993; NID:G3640; PIDN:CAA39465.1; PID:G3641  
 A:Experimental source: strain DBY939  
 R:Chang, T.H.; Arenas, J.; Abelson, J.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 1571-1575, 1990  
 A:Title: Identification of five putative yeast RNA helicase genes.  
 A:Reference number: A34848; MUID:90160368; PMID:2406722  
 A:Accession: A34848  
 A:Molecule type: DNA  
 A:Residues: 316-495, 'I', 497-500 <CHA>  
 C:Genetics:  
 A:Gene: SGD:DBP1  
 A:Cross-references: SGD:S0006040; MIPS:YPL119C  
 A:Map position: 16L  
 C:Superfamily: ATP-dependent RNA helicase DBP1  
 C:Keywords: ATP; nucleotide binding; P-loop  
 F:198-205/Region: nucleotide-binding motif A (P-loop)  
 F:314-319/Region: nucleotide-binding motif B  
 F:318-321/Region: DEAD motif

## Alignment Scores:

Score:	1.71e-71	Length:	617
Percent Similarity:	1108.00	Matches:	252
Best Local Similarity:	57.43%	Conservative:	92
Query Match:	42.07%	Mismatches:	185
	28.02%	Indels:	70
		Gaps:	14

US-09-714-865-15 (1-2172) x S62003 (1-617)

QY 409 AATATATTCAGAGCTTCAGGCGCCATACAGAAAGTGAAGAGTATTC----- 459  
 Db 14 AsnAsnLysGlnuAnGly-----GlyGlyLysSerSerIrrpValProPro 30  
 QY 460 -----CGAGTTCCTGAGAGATTGGCTTACGAAGTCCAAATATGACTTA 507  
 Db 31 HieLysArgSerArgIrrpLysProSerPheGlnArgSerThrProLysGlnGlu--- 48  
 QY 508 GACCCAGCAAGATATATGACGACGACGAGCTTGGCTTTGGTTTACAAACCGATATTA 567  
 Db 49 -----AspLysVal-----ThrGlyLysPhePheArgG--- 59  
 QY 568 AGTGGC-----ACAGTAATGTGATATCTTCAAGACAGAAAGTGCAGTGAAGTGA 621  
 Db 60 AlGlyArgGlnThrGlyAsn----- 66  
 QY 622 CGAGGTGTTACAAAGGTTTAAATGAAGATTAACAGGCTTGAAGAAATTTCTGG 681

Db 67 AsnGlyLysPhePheGlyPheSerLysGlnArgAsnGlyGlyThrSerAlaAsnTyraSn 86  
 QY 682 AAGTCAGAACAGAGAGAGAGAAAGTATGATATCTACAGACCAAAAGTACCTACATA 741  
 Db 87 ArgGlyLysSerSerAsnTyraLysSerSerGlyAsnArgIrrpValAsnGlyLysHisIle 106  
 QY 742 CCC---CTTCCTCCACCTGAGATGAGAGACTCCATCTTTGCA----- 780  
 Db 107 ProGlyProLysAsnAlaLysLysGlnuAlaGlnLysPheGlyValHisAspAspProAsp 126  
 QY 781 CATTATACAGACAGCATTAATCTTCGACAAATACGACACTTCTTGGAGAGTCTGGA 840  
 Db 127 TyHisSerSerGlyLysLysPheAspAsnTyraAspAsnLysProValAspAlaSerGly 146  
 QY 841 CATGATCACCACAGACAGATTCGACTTGTGAAGAAAGCTAATCTCTGACAGACTAAT 900  
 Db 147 LysAspValProGluProIleLeuAspPheSerSerProProLeuAspGlnuLeuMet 166  
 QY 901 AACACATTGCTTAAAGCTGTTATCTAATCTTACCTCTGCAAAAATACAGATTCT 960  
 Db 167 GlnAsnLysLysLeuLysSerPheThrLysProThrProValGlnuLysTyrrSerIlePro 186  
 QY 961 ATCATCTTCAGACAGACAGATTGATGCTTGTCTCAACAGGCTTGGAGACATCGC 1020  
 Db 187 IleValThrLysGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrGly 206  
 QY 1021 GCTTTCCTCCACCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGAT 1068  
 Db 207 GlyPheLeuPheProLeuPheThrGlnuLeuPheArgSerGlyProSerProValProGlu 226  
 QY 1069 -----GCCAGTGGTTTAAAGATTGACAGAACAGAGTGTATTTATGACACCAACT 1122  
 Db 227 LysAlaGlnSerPheTyrrSerArgLysGlyTyrrProSerAlaLeuValAlaProThr 246  
 QY 1123 CGAGATTTGTCACACAGATTTATTTGGAAGCCAGAAATTTCTTTGGACTTGTGTA 1182  
 Db 247 ArgGlnuLeuAlaThrGlnLysPheGlnuAlaArgLysPheThrTyrrArgSerIrrpVal 266  
 QY 1183 AGAGCTGTTTATATATGAGGAGACCCAGCTGGACATTCATTCAGCAAAATGTCACAA 1242  
 Db 267 ArgProCysValValLysGlyAlaProIleGlyAsnLysMetCysGlnuValAspArg 286  
 QY 1243 GGCTGTAATATATATGCTACTCTCTGAGAGACTGATGATATGACAGCAAAAGAA 1302  
 Db 287 GlyCysAspLeuLeuValAlaThrProGlyArgLeuAsnAspLeuGlnuArgGlyLys 306  
 QY 1303 ATTGCTCTCAACAGATCAAACTATGATTTGGATGAGATGATGATGATGATGATGATG 1362  
 Db 307 ValSerLeuAlaAsnLysLysTyrrLeuValLeuAspGlnuAlaAspMetLeuAspMet 326  
 QY 1363 GGTTCGTCACAAATGAAAGATTAATTTCTTCCACAGAAATGCCATCAAAAGACAG 1422  
 Db 327 GlyPheGlnProGlnIleArgHisIleValGlnGluCysAspMetProSerValGlnuAsn 346  
 QY 1423 CGCCAAACCTTATATGTCAGTGCACATTTCCAGAGAAATTTGAAAGTGTGTCAGAG 1482  
 Db 347 ArgGlnThrLeuMetPheSerAlaThrPheProValAspLysGlnHisLeuAlaAspArg 366  
 QY 1483 TTTTAAAGTCAAAATATCTGTTTGTCTGTTGACACAGTGGTGGAGCATGTAGAT 1542  
 Db 367 PheLeu---AspAsnTyrrLysPheLeuSerValGlyValGlySerThrSerGlnuAsn 385  
 QY 1543 GTTCAGAGACGCTTCTCCAGATTTGGCCAGTCTCAAAAAGAGAAAAGCTCGTTGAAT 1602  
 Db 386 IleThrGlnArgLysLeuTyrrValAspAspMetAspLysLysSerAlaLeuLeuAspLeu 405  
 QY 1603 CTCGAAACATAGAGGATGAAGAAAGATATGCTTTGTTGAATGAAGAAAGACGAT 1662  
 Db 406 LeuSerAlaGlnHisLysGlyLeuThrIleuLysPheValGlnuThrLysArgMetAlaAsp 425  
 QY 1663 TTTACTGCAACTTTCTTGTCTCAAGAAAATATCAACTACAGATGATGATGATGATG 1722  
 Db 426 GlnLeuThrAspPheLeuLysMetGlnAsnPheLysAlaThrAlaIleHisGlyAspArg 445







```

Db      373 PheLeuAlaValGlyArgValIglySerSerThrAspLeuLeuValGlnArgValGluPhe 392
Qy      1564 GTTGGCCAGTTCCAAAAAGAAAGAAAGCTGTTGAATTCG-----CGAAACATA 1614
Db      393 ValLeuAspSerAspLysArgSerHisLeuMetCaspLeuLeuHisAlaGlnArgGluAsn 412
Qy      1615 GGG-----GATGAAGAACTATGCTCTTTGTTGAACTAAGAAAAAGCAGAT 1662
Db      413 GlyIleGlnGlyLysGlnAlaLeuThrLeuValPheValGlnThrLysArgGlyAlaAsp 432
Qy      1663 TTTAACGCAACTTTCTTTGTTGCAAGAAAAATATCAACTCAAGTATCATGCTGATCGG 1722
Db      433 SerLeuGluAsnTrpLeuCysIleAsnGlyPheProAlaThrSerIleHisGlyAspArg 452
Qy      1723 GAACAGAGAGAGGGGAGCAAGCTTTCGAGATTTTCGCTTTGAAAGTCCAGTCTT 1782
Db      453 ThrGlnGlnIleuTrgGlnValAlaLeuLysAlaPheLysSerGlyArgThrProIleLeu 472
Qy      1783 GTTGCTACTTCAGTACGCTGCAGAGGCGTGATATTGAAATGCAACATGTTATCAAT 1842
Db      473 ValAlaThrAspValAlaAlaArgGlyLeuAspIleProHisValAlaHisValAlaAsn 492
Qy      1843 TTTCATCTTCCTCTACCATTTGATGAATATGTCATCGAATGGCGCTACTGCTGTGT 1902
Db      493 PheAspLeuProAsnAspIleAspAspTyrValHisArgIleGlyArgThrGlyArgAla 512
Qy      1903 GGGAAATCTGGCAGAGCAATTCCTTTTTCATCTTGAATCCGATACCATTTAGCAG 1962
Db      513 GlyLysSerIleuAlaThrAlaPhePheAsn--AspGlyAsnThrSerLeuAlaArg 531
Qy      1963 CCTCTAGTAAGATATTGACAGATGCTCAACAGAGATGCTTCGTCATGTTGAAAGAAAT 2022
Db      532 ProLeuAlaGlnLeuMetGlnGlnAlaAsnGlnGlnValProGlnTrpLeuThrArgTyr 551
Qy      2023 GCCCTTAGTACATACATTCCTCGCTTCAGTGTAGTACAAAGAGAAACGTGTTGCATCA 2082
Db      552 AlaSerArgSerSerPheGlyGlyGlyLysAsnArgArgSerGlyArgPheGlyGly 571
Qy      2083 GTTGATACCGAGAAAGGCGAGACACTTGAACACACTGGTTTCTTCTACGCA 2139
Db      572 ArgAspPheArgArgGlyGlySer-----PheGlySerGlyArg 584

RESULT 14
T45677
ATP-dependent RNA helicase-like protein - Arabidopsis thaliana
N/Alternate names: protein F14P22.160
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Nov-2000
C/Accession: T45677
R/D:Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23011
A/Accession: T45677
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-646 <DAN>
A/Cross-references: EMBL:AL137082
A/Experimental source: cultivar Columbia; BAC clone F14P22
C/Genetics:
A/Map position: 3
A/Intons: 239/3; 267/3; 348/3; 404/3; 442/3
A/Note: F14P22.160
C/Superfamily: ATP-dependent RNA helicase DBPI

Alignment Scores:
Pred. No.: 1.37e-70 Length: 646
Score: 1095.50 Matches: 258
Best Local Similarity: 57.33% Conservative: 90
Best Local Similarity: 42.50% Mismatches: 194
Query Match: 27.71% Indels: 65
DB: 2 Gaps: 14

```

```

US-09-714-865-15 (1-2172) x T45677 (1-646)
Qy      337 TCTAGTATGATACGCGAAGTATATCCAGACGAGACGAGCTTTTCCAGAGAGC--- 393
Db      2 SerSerAsnSerTrpAlaAspValSerGluSerGlnArgAlaProSerGlyGlyTyr 21
Qy      394 GGCCTATCGAGATGGAATAATATTCAGAGCTTCAGGCGCATACAGAAAGGTGAAGAGT 453
Db      22 GlyTyrSerArgProSerArgThrAsnTyrValProProHisLeuThrSer----- 38
Qy      454 AGTTCCGAGCTTCCGTCGAGAGATTTGGCTCTGAGAGTCCA--AATATGACTTAC 510
Db      39 -----ArgThrProSerSerGluPheValAlaProSerProGlnLysAsnAsp----- 54
Qy      511 CCAGACGAATGTATGACGGCCTGTCGCTTTTGGTCTTGAAGACACGATTAAGT 570
Db      55 -----ArgGlyGlyTyrGlyGly-----Ala 61
Qy      571 GGCACAGTAATGATGATCTTCTCAAGCAGAAAGTGGCAGTGAAGTGAACGAGTGT 630
Db      62 AsnSerGlyTyrGlyArgGlyGlnGlyTyrGlyGlyArgGlySerGlyTyrGlyGly 81
Qy      631 TACAAAGCTTTAATGAGAGATTAATACAGGCTTCGAAAAGATTTCTGGAAGTCA 690
Db      82 -----ArgGlyGlyProValGlyLysTrpAsnAlaArg 92
Qy      691 GCAGAGGAGGAGAAAGTATGATCTCA----- 720
Db      93 SerGlyGlyTyrAspArgArgAspThrGlnThrAsnProPheGlyAsnAspGlyAsnAla 112
Qy      721 GGAACAAAGTGAACCTACATACCCCTCTCCACCTGAGAGATGAGACCTCATCTTTGCA 780
Db      113 AspProAlaValAsn-----GlnGlnGlnAsnThrVal----- 123
Qy      781 CATTATCAGACAGCATTAATCTTCGACAAATACGACATCTTCTTGTGCAAGTGTGCA 840
Db      124 -----IleAsnPheGlnAlaTyrGlnAspIleProIleGlnThrSerGly 138
Qy      841 CATGATGACCCACACCAATTCGACTTTTGAAGAGCTAATCTGTCGACACGTAAT 900
Db      139 AspAsnValProProProValAsnThrPheAlaGlnIleAspLeuGlyGlnAlaLeuAsn 158
Qy      901 AACCAATTCGTAACACTGGTTATTAAGCTTACTCTGTCGCAAAATACGATATTCT 960
Db      159 LeuAsnIleGlnArgLysTyrValLysProThrProValGlnArgAsnAlaIlePro 178
Qy      961 ATCATACTTCGACGACGAGATTTGATGCTTGCTCAAAACAGGCTGGGAAGCTGCG 1020
Db      179 IleLeuAlaIleGlyArgAspLeuMetAlaCysAlaGlnThrGlySerGlyLysThrAla 198
Qy      1021 GCTTTCTCTACCATTTTGGCTCATATGATGCATGAT---GGAATTACTGCCAGTCT 1077
Db      199 AlaPheCysPheProIleIleSerGlyIleMetLysAspIleHisIleGlnArgProArg 218
Qy      1078 TTAAAGAGTTGACAGAACACGAGTGTATTTATTTGAGCAACCACTGAGATTTGCAAC 1137
Db      219 GlyValArgGlyValLysTrpLeuAlaValIleLeuSerProThrArgGlnLeuAlaCys 238
Qy      1138 CAGATTATTTGAGACCCAGAAAATTTCTTTTGGACCTTGTTGAAGCTGTTGTTATA 1197
Db      239 GlnIleHisAspIleAlaArgLysPheSerTyrGlnThrGlyValLysValAlaIle 258
Qy      1198 TATGGGGGAACCCAGCTGGGACATTTCAATTTGCAAAATAGCTAAGGCTTAATATTA 1257
Db      259 TyrGlyGlyThrProValAsnGlnGlnIleArgGlnLeuGlnArgGlyValAlaPheLeu 278
Qy      1258 TGTGCTACTCTCGAAGACTGATGATATCATGCAAGAAAGAAAGTGGTCTCAACAG 1317
Db      279 ValAlaThrProGlyArgLeuAsnAspLeuLeuGlnArgGlyArgValSerLeuGlnMet 298
Qy      1318 ATCAAAATCTAGTTTGGATGAAGCTGATGCATGTTGATATGGCTTTGCTCCAGAA 1377
Db      299 ValArgPheLeuAlaLeuAspGlnAlaAspArgMetLeuAspMetGlyPheGlnProGln 318

```



Db 287 PheAlaGluAlaAsnLeuThrGluThrMetGlnLysAsnValAlaHisAlaGlyTyrSer 306  
 QY 928 AAGCTTACTCTGTGCAAAAATAACAGTATCTTATCATCTTCCAGACAGCATTTGATG 987  
 Db 307 LysThrThrProIleGlnGlnTyrAlaLeuProLeuValHisGlnGlnTyrAspIleMet 326  
 QY 988 GCTTGTGCTCAACACAGGCTTGGGAAAGCTGGCGCTTTCTTCCATCAATTTGGCTCAT 1047  
 Db 327 AlaCysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuSerIleMetThrArg 346  
 QY 1048 ATGATGATGATGATGAATTAACCTGCGAGCTTTAAAGAGTTGACAGACAGAGCTGAT 1107  
 Db 347 LeuIleAspSerAsnLeuAsnThrAlaGlyGluGlyCysTyrProArgCysIle 366  
 QY 1108 ATGTAGCAACCACTGAGATTTGTCACACAGATTTATTTGGAAGCCAGAAAATTTCT 1167  
 Db 367 IleLeuThrProThrArgGluLeuAlaAspGlnIleTyrAsnGluGlyArgLysPheAla 386  
 QY 1168 TTTCGACTTGTGTAAAGCTGTGTATATATATGGGGAAACCAAGCTGGACATTCAT 1227  
 Db 387 TyrGlnThrMetMetGluIleLysProValTyrGlyLeuAlaValGlyTyrAsnLys 406  
 QY 1228 CGACAAATGTACAGGCTGTATATATATATGTGCTACTCTCCAGAGACCTGATGATC 1287  
 Db 407 GlyGlnIleGluLysGlyAlaThrIleLeuAlaGlyThrValGlyArgIleLysHisPhe 426  
 QY 1288 ATAGCAAAAGAAAGATTGCTCAACAGATCAAACTTAAGTTTGTGATGAAGCTGAT 1347  
 Db 427 CysGluGluGlyThrIleLysLeuAspLysCysArgPheValLeuAspGluAlaAsp 446  
 QY 1348 CGCATGTTGAT--ATGGTTTGGTCCAGAAATGAAAGATTATTTCTTGCCCAAGA 1404  
 Db 447 ArgMetIleAspAlaMetGlyPheGlyThrAspIleGluThrIleValAsnTyrAspSer 466  
 QY 1405 ATGCCATCAAGAGACAGCGCCAAACCTATGTTCAGTGCACACTTTCCAGAGCAAT 1464  
 Db 467 MetProArgLysGlnAsnArgGlnThrLeuMetPheSerAlaThrPheProAspSerVal 486  
 QY 1465 CAAGGTGGCTGCAGAGTTTAAAGTCAAAATTATCTGTTGTGCTGTGGAACAAGT 1524  
 Db 487 GlnGluAlaAlaArgAlaPheLeuArgGluAsnTyrValMetIleAlaIleAspLysIle 506  
 QY 1525 GGTGAGCATGTAGATGTTTCAGACAGCCGTTCTCCAGTTGGCCAGTTTCAAAAAGA 1584  
 Db 507 GlyAlaAlaAsnLysCysValLeuGlnGluPheGluArgCysGluArgSerGluLys 526  
 QY 1585 GAAAGCTCGTGAATTCG-----CGA 1608  
 Db 527 AspLysLeuLeuGlnLeuGlyIleAspIleAspSerTyrThrThrGluLysSerAla 546  
 QY 1609 AACATAGGGGATGAAGAACTATGCTTGTGAAACTAAGAAAAAGCAGATTTTACT 1668  
 Db 547 GluValTyrThrLysLysThrMetValPheValSerGlnArgAlaMetAlaAspIleLeu 566  
 QY 1669 GCAACTTTTCTTGTCAAGAAAAATATCAACTACAGTATCCATGCTGATCGGAAACAG 1728  
 Db 567 AlaSerIleLeuSerSerAlaGlnValProAlaIleThrIleHisGlyAlaArgGluGln 586  
 QY 1729 AGAGAGCGGAGCAAGCTCTTGAAGATTTTCGCTTTGGAAGAGCCAGTTCTTGTGCT 1788  
 Db 587 ArgGluArgSerIleAlaLeuArgGlnPheArgAsnGlySerLysProValIleuIleAla 606  
 QY 1789 ACTTCAGTACTGCCAGAGGCTGATATTTGAATAATGCAACATGTTATCAATTTTAT 1848  
 Db 607 ThrAlaValAlaGluArgGlyLeuAspIleLysGlyValAspHisValIleAsnTyrAsp 626  
 QY 1849 CTTCTTTTACATGATGAATAATGTTCAATGAAATTTGGGCTACTGCTGTTGGGAAT 1908  
 Db 627 MetProAspAsnIleAspArgTyrIleHisArgIleGlyArgThrGlyArgValGlyAsn 646  
 QY 1909 ACTGGCAGAGCAATTTCTTTTGTGATCTTGAATCGATTAACCATTTAGCAGACCTCTA 1968  
 Db 647 SerGlyArgAlaThrSerPhe-----IleSerGlnAspCysSerLeuLeuSerGluLeu 664

QY 1969 GTAAAGTATGTACAGATGCTCAACAGAGTGTTCCTGATGTTGGAAGAAATTCCTTT 2028  
 Db 665 ValGlyValLeuAlaAspAlaGlnGlnIleValProAspTyrMetGln----- 680  
 QY 2029 AGTACATACATTCCTGCTTCACTGCTAGTACAGAGAAACGCTTTGTGATCAGTTGAT 2088  
 Db 681 -----GlyAlaAlaGlyLysAsnTyrGlyAlaSer----- 690  
 QY 2089 ACCAGAAAGGCAAGACACTTTGAACACAGCTGGGTTTCTTCTTCACAGAGCTCCCAT 2148  
 Db 691 -----GlyPheGlySerSer---ValProThr 698  
 QY 2149 -----CAAGTATGATGATGATGATG 2169  
 Db 699 GlnValProGlnAspGluGluGlyTyr 707

Search completed: June 10, 2003, 16:44:39  
 Job time: 115.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:42:51 ; Search time 156.5 Seconds  
(without alignments)  
2865.660 Million cell updates/sec

Title: US-09-714-865-15  
Perfect score: 3954  
Sequence: 1 atcg9g9gagatgagatg9gga.....tagatgatgagtcagtcggaat 2172

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 392085 segs, 103240269 residues  
Total number of hits satisfying chosen parameters: 784170

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ n2p model -DEV=x1p  
-Q=/cgn2\_1/USPTO.spool/US09714865/runat\_05062003\_111800\_26140/app\_query.fasta\_1.2311  
-DB=published Applications\_AA -QFMT=fastan -SUFFIX=rabp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.csi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09714865.OCGN.1.1.59 @runat\_05062003\_111800\_26140  
-NGPU=6 -CPU=3 -NO\_MMAP -LARGESOURCE -NBS\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV.TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Published Applications AA:  
1: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	796	20.1	575	9	US-10-108-605-55
2	761	19.2	500	9	US-10-228-897-13
3	709.5	17.9	648	10	US-09-923-831-43
4	704	17.8	1261	12	US-10-147-268-2

5	681.5	17.2	709	9	US-09-736-457-335	Sequence 335, App
6	681.5	17.2	709	9	US-09-902-941-335	Sequence 335, App
7	681.5	17.2	709	9	US-09-849-626-335	Sequence 335, App
8	681.5	17.2	709	9	US-10-017-754-335	Sequence 335, App
9	622	15.7	630	9	US-10-195-117-2	Sequence 2, Appli
10	621	15.7	524	10	US-09-815-242-13491	Sequence 13491, A
11	618.5	15.6	732	9	US-09-738-626-4778	Sequence 4778, Ap
12	614.5	15.5	492	10	US-09-815-242-11308	Sequence 11308, A
13	614.5	15.5	492	10	US-09-732-091-2	Sequence 2, Appli
14	614.5	15.5	504	10	US-09-732-091-42	Sequence 42, Appli
15	601.5	15.2	624	10	US-09-963-790A-2	Sequence 2, Appli
16	599.5	15.2	536	10	US-09-815-242-10774	Sequence 10774, A
17	587	14.8	495	9	US-09-971-536-52	Sequence 52, Appli
18	578	14.6	646	10	US-09-815-242-13769	Sequence 13769, A
19	575	14.5	646	10	US-09-815-242-10311	Sequence 10311, A
20	567	14.3	424	9	US-09-738-626-4359	Sequence 4359, Ap
21	562.5	14.2	613	10	US-09-815-242-10990	Sequence 10990, A
22	561.5	14.2	510	10	US-09-815-242-5256	Sequence 5256, Ap
23	561.5	14.2	526	9	US-10-084-205-74	Sequence 74, Appli
24	561.5	14.2	526	10	US-09-925-637-74	Sequence 74, Appli
25	552.5	14.0	487	10	US-09-815-242-13397	Sequence 12397, A
26	552.5	14.0	487	10	US-09-815-242-12991	Sequence 12991, A
27	552	14.0	567	10	US-09-815-242-11885	Sequence 11885, A
28	545.5	13.8	764	9	US-09-902-627-1	Sequence 1, Appli
29	542	13.7	106	9	US-09-764-891-3857	Sequence 3857, Ap
30	523.5	13.2	459	10	US-09-925-301-1027	Sequence 1027, Ap
31	517	13.1	455	9	US-09-909-567B-48	Sequence 48, Appli
32	515.5	13.0	421	10	US-09-815-242-11703	Sequence 11703, A
33	505.5	12.8	411	10	US-09-755-325-8	Sequence 8, Appli
34	504.5	12.8	415	10	US-09-755-325-2	Sequence 2, Appli
35	503.5	12.7	343	10	US-09-925-301-1439	Sequence 1439, Ap
36	502.5	12.7	448	9	US-09-738-626-5103	Sequence 5103, Ap
37	491.5	12.4	447	10	US-09-815-242-13412	Sequence 13412, A
38	486.5	12.3	447	10	US-09-815-242-13590	Sequence 13590, A
39	486.5	12.3	403	9	US-09-991-496-10	Sequence 10, Appli
40	485	12.3	403	9	US-09-874-923-10	Sequence 10, Appli
41	472.5	11.9	452	10	US-09-815-242-10493	Sequence 10493, A
42	468.5	11.8	466	9	US-10-228-897-12	Sequence 12, Appli
43	467.5	11.8	506	10	US-09-801-368-90	Sequence 90, Appli
44	463.5	11.7	437	10	US-09-815-242-5233	Sequence 5233, Ap

## ALIGNMENTS

RESULT 1  
US-10-108-605-55  
; Sequence 55, Application US/10108605  
; Patent No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; APPLICANT: Kamdar, Klm  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE  
; FILE REFERENCE: 31133B  
; CURRENT APPLICATION NUMBER: US/10/108, 605  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 09/761,142  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/176,418  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-108-605-55  
Alignment Scores: 2.1e-60 Length: 575  
Pred. No.:

Score: 796.00 Matches: 206  
Percent Similarity: 51.18% Conservative: 98  
Best Local Similarity: 34.68% Mismatches: 212  
Query Match: 20.13% Indels: 78  
DB: 9 Gaps: 17

US-09-714-865-15 (1-2172) x US-10-108-605-55 (1-575)

QY 439 AGAGTGGAGAGAGTACGTTCCAGGT-----TCCCGTGAGAGATTGTCAGGA 489  
DB 13 ArgGlyGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 32  
QY 490 AGTCAATAATGACTTACCCAGAGAGAGATGTAGCCGACCTGGTGGCTTTTGGT 549  
DB 33 Asn-----ArgPheGlyGly----- 37  
QY 550 TCTAAGAACCAATTAAGTGGCAAGGTATGCTGATCTTCAAGCAAGTGGC 609  
DB 38 -----GlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 50  
QY 610 AGTGAAGTGAACGAGGTGTTACAAAGCTTAATGAAGAATTAACAGCTGCGA 669  
DB 51 AlaValGluValArgPheArgPheArgPheArgPheArgPheArgPheArgPhe 70  
QY 670 AAGAATCTTGAAGTCAAGACAGAAAGAGAGAAAGTACTGATCAAGACCAAAA 729  
DB 71 PheGlyArgPhe---ArgArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 85  
QY 730 GTGACCTACATACCCCTCTCCACCTGAGAT-----GAGGAC 768  
DB 86 -----PrometArgProValArgPheSerAsnLeuAlaProPheLeuLys 100  
QY 769 TCCATCTTTGACATTAACAGACAGCATTAACCTCCAAATACGACATCTT--- 825  
DB 101 AsnPheTyrgInGluHisProAsnValAlaAsnArgSerProTyrgInGluAlaArgTyr 120  
QY 826 -----GTGAAGTGTCTGACATGATGACAGCAACAGCAATCTGACTTT 870  
DB 121 ArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 139  
QY 871 GAAGAAGCTAATCTCTGTACAGACAGTGAATTAACAATGCTGTTATACTAAG 930  
DB 140 SerGluValHisLeuProAspTyrValMetLeuGluLeuArgGluGlyTyrLeuAla 159  
QY 931 CTTCTCTCTGTCGCAAAAATACAGATTCCTATCACTTGTGAGAGACAGATTAAGCT 990  
DB 160 ProThrAlaIleGlnAlaGlnGlyTyrProIleAlaMetSerGlySerAsnPheValGly 179  
QY 991 TGTCTCAACAGAGGTCTGGAGAGACTGGCTTTTCTTCAACAAATTTGGCTCATATG 1050  
DB 180 IleAlaIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 199  
QY 1051 ATGATGATGGAATTAACCTGCGAGTCTTTAAAGATTGACGAA-----CCAGAG 1101  
DB 200 AsnAsnGln-----GlnProLeuGlnArgGlyAspGlyProIle 212  
QY 1102 TGTATTATATGAGACCAACTCGAAGTGTCTCAACAGATTTATTGGAAGCCAGAAA 1161  
DB 213 AlaLeuValLeuAlaProThrArgGluLeuAlaGlnGlnIleGlnGlnAlaIleThrGlu 232  
QY 1162 TTTTCTTTGGAGCTGTGTAGAGCTGTGTATATATGAGGGAACCCAGCTGGACAT 1221  
DB 233 PheGlySerSerSerTyrValArgAsnThrCysValPheGlyGlyAlaProValGlyGly 252  
QY 1222 TCAATTGCAAAATAGTCAAGGCTGTAAATATATATGCTTACTCTCGAAGACTGATG 1281  
DB 253 GlnMetArgAspLeuGlnArgGlyCysGluIleValIleAlaIleThrProGlyArgLeuIle 272  
QY 1282 GATATCATAGCCAAAGAAAGAAATGCTCTCAACAGATCAAAATCTAGTTTGGATGAA 1341  
DB 273 AspPheLeuSerAlaGlySerThrAsnLeuLysCysThrTyrLeuValLeuAspGlu 292  
QY 1342 GCTATCGCATGTTGATATAGGTTTGTGTCAGAAATGAAGATTAATTTCTTGCCA 1401

DB 293 AlaAspArgLeuSerMetGlyPheGluProGlnIleArgGlySerIleValSer----- 310  
QY 1402 GGAATGCCATCAAGAACGAGCCCAACCTTATGTCAGTGCACATTTTCCAGAGAA 1461  
DB 311 -----GlnIleArgProAspArgGlnThrLeuMetCysSerAlaThrProValGlu 328  
QY 1462 ATTCAAAGTTGGCTGACAGATTTTAAAGTCAAAATATATGCTGTTGCTGTGCAAA 1521  
DB 329 ValLeuGlnLeuAlaGluAspPheLeu---GlyAsnTyrIleGlnIleAsnIleGlySer 347  
QY 1522 GTGGGT---GAGACATGTAGAGATGTTACAGACAGCCGTTCTCCAAAGTGGCCAGTTTCA 1578  
DB 348 LeuGluLeuSerAlaAsnHisAsnIleArgGlnValValAspValCysAspGluPheSer 367  
QY 1579 AAAAGAGAAAGCTCTGTTGAATTTCTGCGAAACATAGCGGATGAA----- 1623  
DB 368 LysGluGluLysLeuLysThrLeuLeuSerAspIleTyrAspThrSerGluSerProGly 387  
QY 1624 AGAAGTATGCTCTTGTGAACCTTAAGAAAGACAGATTTTACTGCAACTTTCTTGT 1683  
DB 388 LysIleIleIlePheValGluThrLysArgValAlaAsnLeuValAlaArgPheIleArg 407  
QY 1684 CAAGAAAAAATATCAATCACTACAGATTCATGATCGGAGACAGAGACGCGAGCAA 1743  
DB 408 SerPheGlyValArgCysGlyAlaIleHisGlyAspLysSerGlnSerGluArgAspPhe 427  
QY 1744 GCTCTTGAGATTTTGGCTTTGGAAGTCCGCTTCTTCTTCTTCTTCTTCTTCTTCTT 1803  
DB 428 ValLeuArgGluPheArgSerGlyLysSerAsnIleLeuValAlaThrAspValAlaAla 447  
QY 1804 AGAGGCTGATATGAAATGTGCAACATGATATCAATTTGATGCTCTTCTTCACTT 1863  
DB 448 ArgGlyLeuAspValAspGlyIleLysTyrValIleAsnPheAspTyrProGlnAsnSer 467  
QY 1864 GATGATATGTCATCGAATTTGGCGTACTGTCGTTGTGGGAATACGTGACAGCAAT 1923  
DB 468 GluAspTyrIleHisArgIleGlyArgThrGlyArgSerAsnThrLysGlyThrSerPhe 487  
QY 1924 TCCCTTTTATCTTGAATGATTAACATTTAGCACAGCTTCTAGTAAAGATTTGACA 1983  
DB 488 AlaPhePhe---ThrLysAsnAsnAlaLysGlnAlaLysAlaLeuValAspValLeuArg 506  
QY 1984 GATGCTCAACAGAGTGTCTCGATGTTGGAAGAAATCTTGTATCATCATCTTCT 2043  
DB 507 GlnAlaAsnGlnIleAsnProAlaLeuGlnAsnLeuAlaArgAsnSerArgTyrAsp 526  
QY 2044 GGCTTCAGTGTAGTCAAGA-----GGAACGCTGTTTGCATCAGTT 2085  
DB 527 GlyGlyGlyGlyProSerArgTyrGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 546  
QY 2086 GATACCAAGAAAGGCAAGACACTTGAACACAGCTGGCTT 2127  
DB 547 GlyPheLysLysGly---SerLeuSerAsnGlyArgGlyPhe 559

RESULT 2  
US-10-228-897-13  
Sequence 13, Application US/10228897  
Publication No. US20030092043A1  
GENERAL INFORMATION:  
APPLICANT: Fisher, Paul B.  
APPLICANT: Kang, Dong-Chul  
APPLICANT: GopalKrisnan, Rahul V.  
TITLE OR INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED  
FILE REFERENCE: A34614-A-PCT-USA (070050, 2121)  
CURRENT APPLICATION NUMBER: US/10/228,897  
CURRENT FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: PCT/US01/06960  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 09/515,363  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 25

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 500
; TYPE: PR
; ORGANISM: moligula oculata
US-10-228-897-13

```

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	2.18e-57	500	167	88	125	58
Percent Similarity:	761.00					
Best Local Similarity:	58.22%					
Query Match:	38.13%					
	19.25%					

US-09-714-865-15 (1-2172) x US-10-228-897-13 (1-500)

```

QY 826 GTGGAAGTCTGTGACATGATGACCCACGCAATTTGACTTTGAAGAGCTAATCTC 885
DB 82 ILeThrValaArgGlyHisAsnCysProLysProValLeuAsnPhetyrGluAlaAsnPh
QY 886 TGTACAGACCTGATATACCAATGCTTAAGCTGTTATCTAAGCTTACTCTGTCGA 945
DB 102 ProAlaAsnValIleCysAspValIleAlaArgGlnAsnPhetrGluProAlaIleGln
QY 946 AATATACAGTATCTTATCATATCTGACAGACGAGATTGAGCTGTGCTCAACAGG 1005
DB 122 AlAGlnGlyTrpProValAlaLeuSerGlyLeuAspMetValGlyValAlaGlnThrGly
QY 1006 TCTGGGAAGCTGGCGCTTTCTCTCAACAAATTTGGCTCATATGATGATCAT- 1056
DB 142 SerGlyLysThrLeuSerGlyLeuLeuProAlaIleValHisIleAsnHisGlnProPhe
QY 1057 -GATGGAATTAACCTGCAGTCGTTTAAAGAGTTGACGGAACGAGAGTGT 1104
DB 162 LeuGlnArgGlyAspGly-----ProIleCys 170
QY 1105 ATTATTTAGACACCACTGCAAGATTTGTCACACAGATTTATTTGGAAGCAGAAATTT 1164
DB 171 LeuValLeuAlaProThrArgGlyLeuAlaGlnGlnValAlaIleAlaGlnTyr 190
QY 1165 TCTTTGGAGCTGTGTAAAGCTGTTTATATATATATATATATATATATATATATATAT 1224
DB 191 CysArgIleCysArgLeuLysSerThrCysIleTyrGlyGlyAlaProLysGlyProGln 210
QY 1225 ATTGCAAAATATAGTACAGAGCTGTATATATATATATATATATATATATATATATAT 1284
DB 211 IleArgAspLeuGlnArgGlyValGlnIleCysIleAlaThrProGlyArgLeuIleAsp 230
QY 1285 ATCATAGGCAAAAGAAAGATTTGCTCAACAGATCAAAATCTAGTTTGGATGAAGCT 1344
DB 231 PheLeuGlnCysGlyLysTrpAsnLeuArgArgThrTrpLeuValLeuAspGluAla 250
QY 1345 GATGCAATGTTGATATGAGTTTGTTCACAGAAATGAAGATTAATTTCTTGCCACAGA 1404
DB 251 AspArgMetLeuAspMetGlyPheGluProGlnIleArgLysIleValAsp----- 267
QY 1405 ATGCCATCAAAAGAAACAGCCCAACCTTATGTTACAGTGAATTTTCCAGAGAAATTT 1464
DB 268 ---GlnIleArgProAspArgGlnThrLeuMetTrpSerAlaThrTrpProLysGlyVal 286
QY 1465 CAAGAGTTGGCTGACAGATTTTAAAGCAAAATATATCTGTTTGGCTGTTGGA----- 1518
DB 287 ArgGlnLeuAlaGlnAspPheLeuLys---AspTyrIleHisIleAsnIleGlyAlaLeu 305
QY 1519 -----CAAGTGGGTGAGCAGATGTAGAGATGTTCCAGCAG 1551
DB 306 GluLeuSerAlaAsnHisAsnIleLeuGlnIleValAlaSerValCysHisAspValGlu-- 324
QY 1552 ACCGTTCTCCCAAGTTGGCCAGTCTCAAAAAGAGAAAAGCTGTTGAAATTCCTGCAAAC 1611
DB 325 -----LysAspGlnLysLeuIleArgLeuMetGlnGlu 335

```

```

QY 1612 ATNAGGGATGAA-----AGACTATGCTCTTTGTTGCAACTAAGAAAAAGCAGAT 1662
DB 336 IleMetSerGlnLysGlnLysTrpIleValPheValGlnThrLysAspGlyCysAsp 355
QY 1663 TTATACGCAACTTTCTTGTGTCAAGAAAAATATCAACTACAGTATCCATGGAGATCGG 1722
DB 356 GlnLeuThrArgLysMetArgAspGlyTrpProAlaMetCylIleHisGlyAspLys 375
QY 1723 GAACAGAGAGAGCGGAGCAAGCTCTTGAGATTTTTCGCTTTGAAAGTCCAGATTTCTT 1782
DB 376 SerGlnGlnGlnArgAspTrpValLeuAsnGlnPheLysHisGlyLysAlaProIleLeu 395
QY 1783 GTTGCTACTTCACTAGTACTGCCAGAGCGCTGAGATTTGAAATATGCAACATGTTATCAT 1842
DB 396 IleAlaThrAspValAlaSerArgGlyLeuAspValGlnAspValLysPheValIleAsn 415
QY 1843 TTGATCTCTCTCTCATATGATGATGATATGTTATGTCATGCAATTTGGGCTATGTCGTTGT 1902
DB 416 TyrAspTrpProAsnSerSerGluAspTyrIleHisArgIleGlyArgThrAlaArgSer 435
QY 1903 GGGAAATATGCGCAGACGCAATTTCTTTTGTGATCTTGAAATCGGATTAACATTTAGCAG 1962
DB 436 ThrLysThrGlyThrAlaTyrThrPhePhe-----ThrProAsnAsnIleLysGln 452
QY 1963 -----CCTCTGATAAAAGTATTGACAGATGCTCAACAGATGTT---CCTGCATGCTTG 2013
DB 453 ValSerAspLeuLeuSerValLeuArgGlnAlaAsnGlnAlaIleAsnProLysLeuLeu 472
QY 2014 GAAGAAATTCCTTTACTATCATATTCCTCGCTTCAAGCTGAGTACAAAGAGA 2067
DB 473 GlnLeuValGlu-----AspArgGlySerGlyArgSerArgGly 485

```

## RESULT 3

```

US-09-923-831-43
; Sequence 43, Application US/09923831
; Patent No. US20020115142A1
; GENERAL INFORMATION:
; APPLICANT: Martelange, Val,rie
; APPLICANT: De Smet, Charles
; APPLICANT: Boon-Falloux, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: L0461/7054
; CURRENT APPLICATION NUMBER: US/09/923, 831
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 09/183,706
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 43
; LENGTH: 648
; TYPE: PR
; ORGANISM: H. sapiens
US-09-923-831-43

```

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	7.26e-53	648	160	77	145	13
Percent Similarity:	709.50					
Best Local Similarity:	60.00%					
Query Match:	40.51%					
	17.94%					

US-09-714-865-15 (1-2172) x US-09-923-831-43 (1-648)

```

QY 850 CCACACGCAATTCGACTTTTGAAGAGCTAATCTCTG---CAGACACTGCAATAACAC 906
DB 237 ProAsnProThrCysThrPheAspAspAlaPheGlnCysTyrProGluValMetGluAsn 256
QY 907 ATTCTAAGCTGGTATTACTAAGCTTACTCTGTGCAAAATATCAAGTATCTATCATATA 966
DB 257 IleLysLysAlaGlyPheGlnLysProThrProIleIleSerGlnAlaThrProIleVal 276
QY 967 CTTCGACGACGAGATTTTGATGCTTGTCTCAACAGAGCTGTGGAGAGACTGCGGCTTTT 1026
DB 967 -----

```

```

Db      277 LeuGlnGlyLeuAspLeuLeuGlyValAlaGlnThrGlyThrGlyLeuGlySerTyr 236
Qy      1027 CTCTTACCAATTTGGCTCATATGATGATGATGATTAATGCTGCACTGCTTTAAAGAG 1086
Db      297 LeuMetProGlyPheIleHisLeuVal-----LeuGlnProSerLeuGlyGln 313
Qy      1087 TTGCAGAACACGAGATGATATATGATGACCAATGCGAATTTGGCAACAGATTTAT 1146
Db      314 ArgAsnArgProGlyMetLeuValLeuThrProThrArgLeuAlaLeuGlnValGlu 333
Qy      1147 TTGCAGACCAAGAAATTTCTTTTGGACCTTGCTAAAGCTGTTGTTATATATGCGGGA 1206
Db      334 GlyGlnCysCysLeuTyrSerTyr---LysGlyLeuArgSerValCysValTyrGlyGly 352
Qy      1207 ACCGAGCTGGAGCACTTCATTCGACAAATAGTACAGGCTGATATATATGCTACT 1266
Db      353 GlyAsnArgAspGlnGlnIleGlnGlnLeuLysGlyValAlaPheIleIleAlaThr 372
Qy      1267 CCTGGAAGACTGATGATATCATATGCAAAAGAAATTTGGCTTCAACAGATCAATATAC 1326
Db      373 ProGlyArgLeuAsnAspLeuGlnMetSerAsnPheValAsnLeuLysAsnIleThrTyr 392
Qy      1327 TTAGTTTGGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
Db      393 LeuValLeuAspGlyAlaAspLysMetLeuAspMetGlyPheGlnProGlnIleMetLys 412
Qy      1387 TTATATTTCTGGCCAGAAATGCCATCAAGAGACAGCGCCAAACCTTATGTTAGTGA 1446
Db      413 IleLeu-----LeuAspValArgProAspArgGlnThrValMetThrSerAla 428
Qy      1447 ACTTTCCAGAGAAATTTCAAGAGTTGGCTGCAAGATTTTAAAGTCAAAATATATCTGTT 1506
Db      429 ThrTrpProHisSerValHisArgLeuAlaGlnSerTyrLeuLysGlnPro---MetIle 447
Qy      1507 GTTCTGCTGGACAGATGGGT---GAGCATGTAAGATGTTCAAGACAGCCGTTCTCAA 1563
Db      448 ValTyrValGlyThrLeuAspLeuValAlaValSerSerValLysGlnAsnIleLeuVal 467
Qy      1564 GTTGGCCAGTTCTCAAAAGAGAAAGCTGTTGAAATTTGCGAATCATGAGGAGAT--- 1620
Db      468 ThrThrGlnGlnGlyLysTrpSerHisMetGlnThrPheLeuGlnSerMetSerThr 487
Qy      1621 GAAAGAACTAGTGTCTTTGGAATCTAGAAAGAAAGAGAAAGATTTTACTGCAATTTCTT 1680
Db      488 AspLysValIleValPheValSerArgLysAlaValAlaAspHisLeuSerSerAspLeu 507
Qy      1681 TGTCAAGAAAAATATCACTACATCAAGTATCCATGCTGATCGGAAACAGAGAGACGGGAG 1740
Db      508 IleLeuGlyAsnIleSerValGlnSerLeuHisGlyAspArgGlnArgAspArgGln 527
Qy      1741 CAAGCTCTTGAGATTTTGGCTTGGAAAGTCCAGTCTTGTGCTCACTTCACTGATGCT 1800
Db      528 LysAlaLeuGlnAsnPheLysThrGlyLysValArgIleLeuIleAlaThrAspLeuAla 547
Qy      1801 GCCAGAGGCTGATATGAAATGTCGCAATGTTATCAATTTGATCTTCTTCAACC 1860
Db      548 SerArgGlyLeuAspValHisAspValThrHisValTyrAsnPheAspPheProArgAsn 567
Qy      1861 ATTGATGAATATGTTCAATGATGATGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db      568 IleGlnGlnTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGlyValSer 587
Qy      1921 ATTTCCTTTTGGATCTTGAATCGATACATCAATTTAGACAGCTCTAGTAAAGTATG 1980
Db      588 IleThrThrLeu---ThrArgAsnAspTrpArgValAlaAspGlnLeuIleAsnIleLeu 606
Qy      1981 ACAGATGCTCAACAGAGATGCTCGCATGCTGATGTTGGAAGAAATTTGCC 2025
Db      607 GluArgAlaAsnGlnSerIleProGlnGlnLeuValSerMetAla 621

```

RESULT 4  
US-10-147-268-2  
Sequence 2, Application US/10147268

```

; Patent No. US20020143154A1
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: CIP130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: PP-01513.104/200130.456D3
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-147-268-2

Alignment Scores:
Pred. No.: 2,89e-52 Length: 1261
Score: 704.00 Matches: 227
Percent Similarity: 41.98% Conservative: 113
Best Local Similarity: 28.02% Mismatches: 264
Query Match: 17,80% Indels: 206
DB: Gaps: 26

US-09-714-865-15 (1-2172) x US-10-147-268-2 (1-1261)
Qy      139 GGACTTTCTCGAAGATCATTTTCATGAAAGTGAATTTGCTGCGGCAATTTTGA 198
Db      52 GlyProGlyThrLysArgGlyPheGlyPheGlyGlyPheAlaIle-----SerAlaGly 69
Qy      199 AACAGAGTCTGCTGATGATTAATGACGATAT---ACATCCACA 213
Db      70 LysLysGlnGlnProLysLeuProGlnSerHisSerAlaPheGlyAlaThrSerSer 89
Qy      244 ATGGCTGTTTGA-----GTGGAAAG 267
Db      90 SerSerGlyPheGlyLysSerAlaProProGlnLeuProSerPheTyrLysIleGlySer 109
Qy      268 AGTTTGGAAACAGAGTTTTCACACAGCAGGTTTGAAGATGATGATGCTGCTGTTTC 327
Db      110 LysArgAlaAsnPheAspGlnGlnAsnAlaTyrPheGlnAspGlnGlu----- 125
Qy      328 TGGAGAGATCTAGTAAT-----GACTGCAAGATAATCCACACGG 369
Db      126 ---GlnAspSerSerAsnValAspLeuProTyrIleProAlaGlnAsnSerProThrArg 144
Qy      370 AACAGAGGTTTTCACAGAGAGCGGCTATCGAGATGGAATATTCAGAGCTTCAGGG 429
Db      145 GlnGlnPheHisSerLys----- 150
Qy      430 CCATACAGAGAGGTGAAGAGTATGTTCCGAGTTGCCGTGAGGATTTGGCTTAGGA 489
Db      150 ----- 150
Qy      490 AGTCCAAATATGACTTAGACCCAGACGAATGATGACGCGCACTGGTCTTTTGGT 549
Db      151 ProValAspSerSerAspAspAsp----- 159
Qy      550 TCTAGGAAGACAGATTAATGTCGACAGTATGCTGATCTTCTCAAGACAGA----- 603
Db      160 -----ProLeuGlnAlaPheMetAlaGlnValGlnLysPheGlnAlaAlaAspMet 176
Qy      604 -----AGTGCAGTGGAGTGAACGAGTGGTTTCAAGGTTTAAATGAAGAGTA 654
Db      177 LysArgLeuGlnGlnLysAspLysGlnArgLysAsnValLysGlyIleAspAspPheIle 196
Qy      655 ATACAGGCTCTGGAAGAAATTTCTTG----- 681
Db      197 GlnGlnGlnLysAspGlnGlnAlaTyrPheArgTyrMetAlaGlnAsnProThrAlaGly 216
Qy      682 -----AAGTCAGAGACAGAGAGAGAGAGAAAGTATGATCTCAAGA----- 723
Db      217 ValValGlnGlnGlnGlnGlnLysAsnLeuGlnLysAspSerAspGlyAsnProIleAla 236

```



```

QY 724 CCAAAATGACCTACATACCCCTCTCCACCTGAGAT----- 762
Db 237 ProThrIysLysIleIleAspProLeuProIleAspHisSerGluIleAspTyPro 256
QY 763 -----GAGGACCTCCATCTTGACATTATACAGCAGCATTAAC----- 801
Db 257 ProPheGluLysAsnPheTyAsnGluHisGluGluIleThrAsnLeuThrProGluGln 276
QY 802 ----TTCGACAAATACGACACTATTCTTGGAAGTGTCTGGACACTGATGACACGACGCA 858
Db 277 LeuIleAspLeuArgHisLysLeuAsnLeuArgValSerGlyAlaIleProProArgPro 296
QY 859 ATTCTGACTTTTGAAGAAGCTAATCTCTGTACAGACTGAATTAACAATTCTTAAGCT 918
Db 297 GlySerSerPheAlaHisPheGlyPheAspGluGlnLeuMetHisGlnIleArgLysSer 316
QY 919 GGTTATCTAGCTTACTCTCTGTCGCAAAATACAGTATCTCTATACATCTTGACGACGCA 978
Db 317 GluTyThrGlnProThrProIleGlnCysGlnGlyAlaProValAlaLeuSerGlyArg 336
QY 979 GATTGTAGCTGTGTCTCAACAGGCTGTGGAGAAGCTGGGCTTTCTCTACCAATT 1038
Db 337 AspMetIleGlyIleAlaLysThrGlySerGlyLysThrAlaAlaPheIleTyrProMet 356
QY 1039 TTGGCTCATATGATGATGATGAATAACTGCCAGTGTGTTTAAAGATTGACAGAA--- 1095
Db 357 LeuIleHisIleMetAspGln-----LysGluLeuGluProGly 369
QY 1096 -----CCAGAGCTATTATTGTAGCACCAACTGCAATTTGGTCAACGATTATTG 1149
Db 370 AspGlyProIleAlaValIleValCysProThrArgLysLeuLysCysGlnGlnIleHisPro 389
QY 1150 GAAGCCCAAAATTTCTTTTGGGACTGTGTAGAGCTGTGTATATATATATATATATAT 1209
Db 390 GluCysLysArgPheGlyLysAlaTyAsnLeuArgSerValAlaValIleGlyGlyGly 409
QY 1210 CAGCTGGACATTCATTCGACAAATAGTACAAAGCTGTATATATATATATATATAT 1269
Db 410 SerMetThrGluGlnAlaLysAlaLeuGlnGluGlyAlaGluIleValCysThrPro 429
QY 1270 GGAAGACTGATGATATCATAGGCAAAAGAAAGTGTCTCAACAGATCAATACTTGA 1339
Db 430 GlyArgLeuIleAspHisValLysLysLysAlaThrAsnLeuGlnAspValSerTyLeu 449
QY 1330 GTTTTGTAGAGCTGATCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
Db 450 ValPheAspLeuAlaAspArgMetPheAspMetGlyPheGluTyGlnValArgSerIle 469
QY 1390 ATTTCTTGCCCAAGGATGCCATCAAGGAACAGGCCCAACCCCTTATGTTCACTGCACT 1449
Db 470 AlAsp-----HisValArgProAspArgGlnThrLeuLeuPheSerAlaThr 485
QY 1450 TTTCGAGAGCAATTCAAAGTTGGCTGCAGAGTTTAAAGTAAATTAATCTGTTGTT 1509
Db 486 PheArgLysLysIleGluLysLeuAlaArgAspIleLeu---IleAspProIleArgVal 1504
QY 1510 GCTGTGGACAAGGGGTGGAGCATGTAGAGATGTTCAGACAGCCGT-----CTGCA 1563
Db 505 ValGlnGlyAspIleGlyAlaAsnGluAspAlaThrGlnIleValGluIleLeuHis 524
QY 1564 GTTGGCCAGTTCTCAAA-----AGAGAAAGCTCGTTGAAATTCCTGCGAAG 1611
Db 525 SerGlyPro---SerLysTyrAsnThrLeuThrArgArgLeuValGluPhe----- 540
QY 1612 ATGAGGGATGAAGAACTATGCTTTGTTGAACCTAAGAAAGAAAGCAAGATTTACTGCA 1671
Db 541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnAlaGluGluLeuAla 560
QY 1672 ACTTTTCTTTGTCAAGAAAAAATATCAACTACAGATGCTGATGCTGGGACAGAGAGA 1731
Db 561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580
QY 1732 GAGCGGAGCAAGCTCTTGAGATTTTCGCTTGGAAAGTCCAGTCTTGTGTTGCTACT 1791

```

```

Db 581 GluArgAsnLysValIleSerAspPheLysLysAspIleProValLeuValAlaThr 600
QY 1792 TCAGTGGCTCCGAGAGGCTGATATTTGAAAGTGCACAACTGTTATTCATTTGATCTT 1851
Db 601 AspValAlaAlaArgGlyLeuAspIleProSerIleLysThrValIleAsnTyAspVal 620
QY 1852 CCTTCTACCATTTGATGATATGTTTCATGCAATTTGGCGGATCTGCTGTTGGGAATACT 1911
Db 621 AlaArgAspIleAspThrHisThrHisArgIleGlyArgThrGlyArgAlaGlyLys 640
QY 1912 GCGAGAGCAATTTCTTTTGTGATCTTGAATCGGATACCATTTAGCACAGCTCTAGTA 1971
Db 641 GlyValAlaTyThrLeuLeu---ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659
QY 1972 AAGGATTTGACAGATGCTCAACAGAGATGTTCTC----- 2004
Db 660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGluLeuLeuAspLeuAlaMetGln 679
QY 2005 ---GCATGTGTGGAGAAATTTGCTTTAGT----- 2031
Db 680 AsnAlaThrPheArgLysSerArgPheLysGlyLysGlyLysLysLeuAsnIleGly 699
QY 2032 -----ACATACATTTCTGCTTCACTGATGATGATGATGATGATGATGATGATGAT 2067
Db 700 GlyGlyGlyLeuGlyTyArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719
QY 2068 AAC----- 2070
Db 720 AsnAsnAsnValMetSerAsnTyGluAlaTyLysProSerThrGlyAlaMetGlyAsp 739
QY 2071 -----GTGTTTGA 2079
Db 740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTyLysSerHisPheValAlaAla 759
QY 2080 TCAGTTCATTCACGAAG---GCCAAGACCATTTGACACAGCTGGCTTTCTTCTTCA 2136
Db 760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTyrThrSerAla 779
QY 2137 CGAGCTCCCAATCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2166
Db 780 GlySerLeuAsnSerValProThrAsnSer 789

RESULT 5
US-09-736-457-335
: Sequence 335, Application US/09736457
: Patent No. US20020168637A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Tongtong
: APPLICANT: Bangur, Chaltanya S.
: APPLICANT: Lodes, Michael A.
: APPLICANT: Fanger, Gary
: APPLICANT: Vedvick, Tom
: APPLICANT: Carter, Darick
: APPLICANT: Retter, Marc
: APPLICANT: Mannion, Jane
: APPLICANT: Fan, Liqun
: APPLICANT: Wang, Aijun
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.478C15
: CURRENT APPLICATION NUMBER: US/09/736,457
: CURRENT FILING DATE: 2000-12-13
: NUMBER OF SEQ ID NOS: 1864
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 335
: LENGTH: 709
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-736-457-335
Alignment Scores: 2.05e-50 Length: 709
Pred. No.:

```

Score:	681.50	Matches:	195
Percent Similarity:	45.98%	Conservative:	91
Best Local Similarity:	31.35%	Mismatches:	213
Query Match:	17.24%	Indels:	123
DB:	9	Gaps:	18
US-09-714-865-15 (1-2172) x US-09-736-457-335 (1-709)			
QY 619	GAACGAGTGGTTTACAAAGTTTAAATGAGAAAGTAATTAACAGGCTCTGGAAAGAAATTC	678	
Db 21	GIuArGlyAsnVallyGslYlleaRgAspApriIleGIuGIuLAspAspGIuLAla	40	
QY 679	TGG-----AACTCAGAGCGAAGCA	699	
Db 41	TyrPheArGTYrMeAlAGluAsnProThraGIyAlValGIuGIuGIuGIuAsp	60	
QY 700	GGAGAAAGTAGTACATCTCAAGA-----CCAAAAGTACCTACATACCCCT	747	
Db 61	AsnLeuGIuTYrAspSerAspGIyAsnProIleAlaProThrLyGIleIleAspPro	80	
QY 748	CCTCCAGCTGAGAT-----GAGAGCTGCATCTTTGCA	780	
Db 81	LeuProProlleAspRHiserGIuIleAspTYrProPheGIuLyAsnPheTYrAsn	100	
QY 781	CATTATCAGACAGCATTAAC-----TTGCAGAAATAGACACTAT	822	
Db 101	GIuHIsGIuGIuIleThrAsnLeuThrProGIuGIuLeuIleAspLeuArGHisLyLeu	120	
QY 823	CTTGAGAAAGTGTCTGACATGATGACACACAGCAATTCGACTTTGAGAAAGCTAAT	882	
Db 121	AsnLeuAlaGValSerGIyAlaAlaArProTroArProGIySerSerPheAlaHisPheLy	140	
QY 883	CTGTGTGACACATGAATTAACAACTTGTCTAAAGTGTATACTAACTTACTCTGTG	942	
Db 141	PheAspGIuGIuLeuMeHIsGIuIleArGlySerGIuTYrThrGIuProThrProIle	160	
QY 943	CAAAAATACAGTATTCCTATCATCTACTTGACAGACAGCAATTTGATGGCTTGCTCAACA	1002	
Db 161	GIuCYsGIuGIuValProValAlaLeuSerGIyAlaGAspMetIleGIyIleAlaLyThr	180	
QY 1003	GGGTGTGGAAAGACGCGGCTTTTCTCCACCAATTTGGCTCATATGATGCATGATGA	1062	
Db 181	GIySerGIyLySThrAlaIlePheIleTrpMetLeuIleHisIleMeAspGIu---	199	
QY 1063	ATTACTGCCAGTCTTTTAAAGATTGACGAA-----CCAGAGTGTATTATTGTA	1113	
Db 200	-----LyGIuLeuGIuProGIyAspGIyProIleAlaValIleAla	213	
QY 1114	GCACCAATCTCAGAAATTTGTCACACAGATTTATTGGAAGCCAGAAAATTTCTTTTGGG	1173	
Db 214	CysProThrArGSluLeuCYsGIuGIuIleHisAlaGIuCYsArGpHeGIyLyAla	233	
QY 1174	ACTGTGTAAAGAGCTGTGTATTATATGAGGGGGAACCCAGCTGGGACATTCAATTTGACAA	1233	
Db 224	TyrAsnLeuArGSerValAlaValIyGIyGIyGIySerMetTrpGIuGIuAlaLyAla	253	
QY 1224	ATAGTACAGAGCTGTATATATATTATGTCTACTCTCTGGAAGACTGATATACATGAC	1293	
Db 254	LeuGIuGIuGIuAlaGIuIleValIyAlaCYsThrProGIyArGSluIleAspHisValLyS	273	
QY 1294	AAAGAAAAGATTGGTCTCAACAGATCAAAATCTTAGTTTGGATGAAGCTGATGCGATG	1353	
Db 274	LySlySAlaIthrAsnLeuGIuArGValIserTYrLeuValIlePheAspGIuAlaAspArGMet	293	
QY 1354	TTGATATAGGGTTTGGTCCGAAGTAAGAAAGTAATTTCTTGCCAGAAATGCCATCA	1413	
Db 294	PheAspMetGIyPheGIuTYrGIuValArGSerIleAlaSer-----HisVal	309	
QY 1414	AAGGAACAGCGCCAAACCTTATTTCTAGTGCAGCACTTTCCAGAGAAATTCAAAGTTG	1473	
Db 310	ArGProAspArGInThrLeuIleuPheSerAlaThrPheArGlyLyGIleGIuLySleu	329	
QY 1474	GCTGCAGAGTTTAAAGTCAAATTAATCTGTTGTGCTGTGGACAAAGTGGGTGGACCA	1533	

[illegible]

APPLICANT: Marnetakis, Margarita  
 APPLICANT: Carter, Darick  
 APPLICANT: Fanger, Gary R.  
 APPLICANT: Vedrick, Thomas S.  
 APPLICANT: Bangur, Chaitanya S.  
 APPLICANT: McNabb, Andria  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 FILE REFERENCE: 210121.478C17  
 CURRENT APPLICATION NUMBER: US/09/902,941  
 CURRENT FILING DATE: 2001-07-10  
 NUMBER OF SEQ ID NOS: 2002  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 335  
 LENGTH: 709  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-902-941-335

Alignment Scores:  
 Pred. No.: 2,05e-50 Length: 709  
 Score: 681.50 Matches: 195  
 Percent Similarity: 45.98% Conservative: 91  
 Best Local Similarity: 31.35% Mismatches: 213  
 Query Match: 17.24% Indels: 123  
 Gaps: 18

US-09-714-865-15 (1-2172) x US-09-902-941-335 (1-709)

QY 619 GAACGAGGTGTTAAAGTTTAAATGAGACTAATTAACAGGCTCTGAAAGATTCT 678  
 Db 21 GlnAtrGysAsnValysGlyLeaRgAspRlleGluGluAspGlnGluA 40  
 QY 679 TGG-----AAGTCAGAAAGCAGAGGA 699  
 Db 41 TyrPheArGlyrMetAlaGluAsnProThraGlyValaGlnGluGluGluAsp 60  
 QY 700 GGAGAAAGTAGTACTACTCAAGGA-----CCAAAGTAGCTTACATACCCCT 747  
 Db 61 AsnLeuGluYrAspSerAspGlyAsnProIleAlaProThrIrylsylleIleAspPro 80  
 QY 748 CTCCTCACCCTGAGAT-----GAGACTCCATCTTTGCA 780  
 Db 81 LeuProIleAspRlleAspRlleAspRlyrProProPheGlnIlyAsnPhetyrAsn 100  
 QY 781 CATTATGACAGCGCATTAAC-----TTCCGAAATACGACACTATT 822  
 Db 101 GlnHleGluGluLeuThraAsnLeuThrProGlnGlnLeuIleAspLeuArgHslyLeu 120  
 QY 823 CTGTGGAAGTGTCTGACATGATGACACACAGCAATTTGCACTTTGAAGAGCTAAT 882  
 Db 121 AsnLeuArgValSerGlyAlaAlaProProArgProGlySerSerPheAlaHsPheGly 140  
 QY 883 CTCTGCAGACACTGAATTAACACATTTGTAAGCTGTTACTTACTTCTGTG 942  
 Db 141 PheAspRlleGluGlnMetHsGlnIleArgysSerGluYrThrGlnProThrProIle 160  
 QY 943 CAAATAATACAGTATCTTATCATCTTGAAGAGAGATTTGAGGCTTGCTCAACA 1002  
 Db 161 GlnCysGlnGlyAlaProValAlaLeuSerGlyrAspMetIleGlyIleAlaLysThr 180  
 QY 1003 GGGTCTGGAGAGACTGGGCTTTCTCTACCAATTTGGCTCATATGATGATGAGGA 1062  
 Db 181 GlySerGlyYrThrAlaAlaPheIleTrpPrometLeuIleHsIleMetAspGln--- 199  
 QY 1063 ATAACTCCAGCTGCTTTAAAGCTGAGAA-----CCAGAGTATTATTGTA 1113  
 Db 200 -----LysGlnLeuGlnProGlyAspGlyProIleAlaValIleVal 213  
 QY 1114 GCACCACTGAGAGATTGTTAAACGATTATTGGAAGCAGAAATTTCTTTGGG 1173  
 Db 214 CysProThraArgIleuGlyCysGlnGlnIleHsIleAlaGlyCysLysArgPheGlyLysAla 233

QY 1174 ACTTGTAAGAGCTGTTGTAATATATGAGGGAACCCAGCTGGACATTTCACAA 1233  
 Db 234 TyrAsnLeuArgSerValAlaValTyrGlyGlySerMetRprGlnAlaLysAla 253  
 QY 1234 ATAGTACAGGCTGTAATATATATATATGCTACTCTCGAAGACTGATGATATATAGGC 1293  
 Db 254 LeuGlnGluIlyAlaGluIleValaCysThrProIlyrArgLeuIleAspHsValLys 273  
 QY 1294 AAAGAAAGATTTGCTCAACAGATCAATTAATCTTATTTGTAAGCTGATGCGATG 1353  
 Db 274 LysLysAlaThrAsnLeuGlnArgValSerTyrLeuValaPheAspGlnAlaAspArgMet 293  
 QY 1354 TTGATATGGGTTTTGCTCCAGAAATGAAAGTTAATTTCTTCCAGAAATCCATCA 1413  
 Db 294 PheAspMetGlyrPheGluYrGlnValaArgSerIleAlaSer-----HisVal 309  
 QY 1414 AAGGAACAGGCCCAACCCCTTATGTTAGTGCACATTTTCCAGAGAAATTCAAAGTTG 1473  
 Db 310 ArgProAspArgGlnThrLeuLeuPheSerAlaThrPheArgLysIleGluLysLeu 329  
 QY 1474 GCTGCAGAGTTTTTAAAGTCAAAATTATCTGTTGCTGTCGTGGACAAGGGGTGAGACA 1533  
 Db 330 AlaArgAspIleLeu---IleAspProIleArgValaGlnGlyAspIleGlyGluAla 348  
 QY 1534 TGTAAGATGTTCCAGACAGCCGTT-----CTCCAAAGTTGCCAGTTCTCAAA 1581  
 Db 349 AsnGluAspValThrGlnIleValaGlnIleLeuHsSerGlyPro---SerLysTrpAsn 367  
 QY 1582 -----AGAGAAAGCTCTTGAAATTTCTCGAAACATAGGGAGTGAAGAACTATGTC 1635  
 Db 368 TrpLeuThraArgArgLeuValaGluPhe-----ThrsSerGlySerValLeuLeu 384  
 QY 1636 TTTGTTGAACCTAAGAAAGAAAGATTTTACTGCAACTTTCTTCTTCAAGAAATA 1695  
 Db 385 PheValThrLysLysAlaAsnAlaGluGluLeuAlaAsnAsnLeuLysGlnGlyHis 404  
 QY 1696 TCAACTACAAATATCCATGCTGATCGGGAACAGAGAGCGGAGCAAGCTCTTGAGAT 1755  
 Db 405 AsnLeuGlyLeuLeuHsGlyAspMetAspGlnSerGlnArgAsnLysValIleSerAsp 424  
 QY 1756 TTTGCTTTGAAAGTCCAGCTCTTGTGCTACTTCACTGAGTGTGCGACAGGCTGAT 1815  
 Db 425 PheLysLysLysAspIleProValaIleAsnValaIleThrAspValaAlaAlaArgGlyLeuAsp 444  
 QY 1816 ATTGAATAATGCAACATGTTATCAATTTGATTTCTTCTTCAATGATGATATTGTT 1875  
 Db 445 IleProSerIleLysThrValIleAsnTrpAspValaAlaArgAspIleAspThrHisThr 464  
 QY 1876 CATCGAATGGGCGTACTGCTGTTGGGAAATACGACAGCAATTTCTTTTGGAT 1935  
 Db 465 HisArgIleGlyrThrGlnArgAlaGlyGluLysGlyValaAlaTyrThrLeuLeu--- 483  
 QY 1936 CTGGAATCGATTAACATTTGAACAGCTCTTAATAAGTATTGACAGATGCTCAACAG 1995  
 Db 484 ThrProLysAspSerAsnPheAlaGlyAspLeuValaArgAsnLeuGlnGlyAlaAsnGln 503  
 QY 1996 GATGTTCT-----GCATGCTTGAAGAAATTGCC 2025  
 Db 504 HisValSerLysGluLeuLeuAspLeuAlaMetGlnAsnAlaTrpPheArgLysSerArg 523  
 QY 2026 TTATAGT-----ACATAC 2037  
 Db 524 PheLysGlyGlyLysGlyLysLysLeuAsnIleGlyGlyGlyLeuGlyTyrArgGlu 543  
 QY 2038 ATTCTGCTTCACTGCTAGTAC-----AGAGGAAC----- 2070  
 Db 544 ArgProGlyLeuGlySerGluAsnMetAspArgGlyAsnAsnAsnValaMetSerAsnTyr 563  
 QY 2070 ----- 2070  
 Db 564 GluAlaTyrLysProSerThrGlyAlaMetGlyAspArgLeuThrAlaMetLysAlaAla 583  
 QY 2071 -----GTGTTTGCATCACTTGATACAGAAAG---GGC 2100

Db 584 PheGlnSerGlnTyrLysSerHisPheValAlaIleSerLysSerAsnGlnLysAlaGly 603  
Qy 2101 AAGGCACTTGAACACAGCTGGTTTCTTCTTACAGAGCTCCCAATCCAGTGATGAT 2160  
Db 604 SerSerAlaAlaGlyAlaSerGlyTyrPheSerAlaGlySerLysLeuAsnSerValProThr 623  
Qy 2161 GAGTCA 2166  
Db 624 AsnSer 625  
RESULT 7  
US-09-849-626-335  
Sequence 335, Application US/09849626  
Publication No. US20020197669A1  
GENERAL INFORMATION:  
APPLICANT: Bangur, Chaitanya  
APPLICANT: Fanger, Gary  
APPLICANT: Wang, Aijun  
APPLICANT: Wang, Tonglong  
APPLICANT: Switzer, Anne  
APPLICANT: McNeill, Patricia  
APPLICANT: Clapper, Jonathan  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C16  
CURRENT APPLICATION NUMBER: US/09/849,626  
CURRENT FILING DATE: 2001-05-03  
NUMBER OF SEQ ID NOS: 1926  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 335  
LENGTH: 709  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-09-849-626-335  
Alignment Scores:  
Pred. No.: 2,05e-50 Length: 709  
Score: 681.50 Matches: 195  
Percent Similarity: 45.98% Conservative: 91  
Best Local Similarity: 31.35% Mismatches: 213  
Query Match: 17.24% Indels: 123  
Gaps: 18  
US-09-714-865-15 (1-2172) x US-09-849-626-335 (1-709)  
Qy 619 GAACGAGGTGCTTCAAAAGTTTAAATGAGAAGTAAATACAGGCTTGAAAGATTCT 678  
Db 21 GlnArgLysAsnValLysGlyIleArgAspAspIleGlnGlnLysAspGlnLysAla 40  
Qy 679 TGC-----AACTCAGAACGCAAGGA 699  
Db 41 TyrPheArgTyrMetAlaGlnAsnProThrAlaGlyValValGlnGlnGlnGlnLysP 60  
Qy 700 GGAGAAAGTACTGACTTACTCAAGA-----CCAAAAGTCACTTACATACCCCT 747  
Db 61 AsnLeuGlnTyrAspSerAspGlyAsnProIleAlaProThrLysLysIleIleAspPro 80  
Qy 748 CTCGACCTGAGGAT-----GAGGACTCCATCTTGGCA 780  
Db 81 LeuProIleAspHisSerGlnLysAspTyrProProPheGlnLysAsnPheTyrAsn 100  
Qy 781 CATATTCAGACAGCATAAAC-----TTGCAACAATAGACACTATT 822  
Db 101 GlnHisGlnGlnLysIleThrAsnLeuThrProGlnGlnLysLeuAspLysArgHisLysLeu 120  
Qy 823 CTGTGGAAGTGTCTGACATGATGACACAGCAATTGCACTTGAAGTGAAGACTAT 882  
Db 121 AsnLeuArgValSerGlyAlaAlaProForArgProGlnSerSerPheAlaHisPheGly 140  
Qy 883 CTCTGTCAAGACACTGAATTAACAACATTGCTAAAGCTGTGTTACTTAAGCTTCTCTGTG 942  
Db 141 PheArgGlnGlnLeuMetHisGlnIleArgLysSerGlnTyrThrGlnProThrProIle 160

Qy 943 CAAAAATACAGTATTCCTATCATCTTGACAGACAGATTTGATGGCTTGCAAAACA 1002  
Db 161 GlnCysGlnLysValProValAlaLeuSerGlyArgAspMetIleGlyIleAlaLysThr 180  
Qy 1003 GGGTCTGGGAAGACTGGGGCTTTTCTCTCAATTTTGGCTCATATGATGATGATGA 1062  
Db 181 GlySerGlnLysThrAlaAlaPheIleTyrProMetLeuIleHisIleMetAspGln--- 199  
Qy 1063 ATAACTGCCAGTGGTTTAAAGACTGGCAGAA-----CCAGAGCTATTATTGTA 1113  
Db 200 -----LysGlnLeuGlnProGlyAspGlyProIleAlaValIleVal 213  
Qy 1114 GCACCACTCGAGATTTGTCACACAGATTTATTTTGAAGCCGAAATTTTCTTTGGG 1173  
Db 214 CysProThrArgLysLeuCysGlnGlnIleHisAlaGlnLysLysArgPheGlyLysAla 233  
Qy 1174 ACTTGTTAAGACTGTTTATATATATGAGGGGAACCCAGCTGGACATTCAATTCACAA 1233  
Db 234 TyrAsnLeuArgSerValAlaValIleGlyGlySerMetTyrGlnAlaLysAla 253  
Qy 1234 ATAGTACAGGCTGTATATATATATGCTACCTCGGAGACTGATGATCATATGCC 1293  
Db 254 LeuGlnGlnLysAlaGlnLysValValCysThrProGlyArgLeuIleAspHisValLys 273  
Qy 1294 AAAGAAAGATTGGCTCAACAGATCAATCAATTAAGTATTTTGGATGAGCTGATGCCATG 1353  
Db 274 LysLysAlaThrAsnLeuGlnArgValSerTyrLeuValPheAspGlnAlaAspArgMet 293  
Qy 1354 TTGGATATGGGTTTGGTCCAGAAATGAAAGCTTAATTTCTTGCCAGGAATGCCATCA 1413  
Db 294 PheAspMetGlyPheGlnTyrGlnValArgSerIleAlaSer-----HisVal 309  
Qy 1414 AAGGAACAGCCCAACCTTATGTTTCAAGTCACTTTTCCAGAGAAATTCAAAGTTG 1473  
Db 310 ArgProAspArgGlnThrLeuLeuPheSerAlaThrPheArgLysLysIleGlnLysLeu 329  
Qy 1474 GCTGCAGACTTTTAAAGTCAAAATTAATCTTGTGCTGTCGTGACAAAGTGGTGAGCA 1533  
Db 330 AlaArgAspIleLeu---IleAspProIleArgValValGlnGlnLysAspIleGlyLysAla 348  
Qy 1534 TGTAAGATGTTCAAGACAGCCGTT-----CTCCAACTGGCCGATTTCTCAAAA----- 1561  
Db 349 AsnGlnAspValThrGlnIleValGlnIleLeuHisSerGlyPro---SerLysTyrAsn 367  
Qy 1582 -----AGAAAAAGCTCTGTAATTCGCAAACTTAGGGATGAAAGATCATAGTGC 1635  
Db 368 TrrPheThrArgArgLeuValGlnPhe-----ThrsSerGlySerValLeu 384  
Qy 1636 TTGTGAAACTAAGAAAGAAAGAGATTTTACTGCACTTTTCTTGTCAAGAAATA 1695  
Db 385 PheValThrLysLysAlaAsnAlaGlnGlnLysLeuAlaAsnLeuLysGlnGlnLysHis 404  
Qy 1696 TCAACTCAAGATATCATGTGATCGGGAACAGAGAGCGGAGCAAGCTTTGAGAT 1755  
Db 405 AsnLeuGlnLeuLeuHisGlyAspMetAspGlnSerGlnLysArgLeuValIleSerAsp 424  
Qy 1756 TTTGCTTTGGAAGATGCCAGTCTTGTGCTTACTTCACTAGTGGCCAGAGGCTGGAT 1815  
Db 425 PheLysLysLysAspIleProValLeuValAlaThrAspValAlaAlaArgGlyLeuAsp 444  
Qy 1816 ATTGAANAATGCAACATGTTATCAATTTGATCTTCTTCACTTCAATGATTAATGTT 1875  
Db 445 IleProSerIleLysTrpValIleAsnTyrAspValAlaArgAspIleAspThrHisThr 464  
Qy 1876 CATCGAATGGCGCTACTGTCGTGTTGGAAATATCGCAGAGCAATTCCTTTTGGAT 1935  
Db 465 HisArgIleGlnArgThrGlnArgAlaGlyGlnLysGlyValAlaTyrThrLeuLeu--- 483  
Qy 1936 CTGAATCGATTAACATTTTGAACAGACCTCTTACTAAAGATTTATGACAGATGCTCAACAG 1995  
Db 484 ThrProLysAspSerAsnPheAlaGlyAspLeuValArgValAsnLeuGlnGlyAlaAsnGln 503





```

Db      240 1GlyIleValIleProThrProIleGlnSerGlnAlaIleProIleIleLeuGlnGlyI 260
Qy      978 AGATTGATGGCTTGCTCAAAACAGGGCTGGAGACAGCGCTTTCTCCACCAAT 1037
Db      260 eaapleuIleValIleGlnThrGlyThrGlySerThrLeuSerThrLeuSerThrProG 280
Qy      1038 TTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1097
Db      280 yPheIleIleLeu-----AspSerGlnProIleSerArg---GluGlnArgsnGlyPr 297
Qy      1098 AGAGTATATATGTAGACCAACTGAGAAATGGTCAACCAAGATTTATTTGGAGGACG 1157
Db      297 ogIleLeuValIleThrProThrArgGlyLeuAlaLeuIleValGlnIleGlyCyse 317
Qy      1158 AAAATTTCTTTGGGACTTGTAAGAGCTTGTTGTAATATATGAGGGGAGCCAGCGGG 1217
Db      317 rlySlySerIle---LysGlyLeuIleSerIleCysIleThrGlyArgsnArgas 336
Qy      1218 ACATTCATTCGACCAATAGTACAGGCTGTAAATATATATGCTTACTCTGGAAGACT 1277
Db      336 ngIyGlnIleGlyAspIleSerIleGlyValAspIleIleAlaIleThrProGlyArgle 356
Qy      1278 GATGATATCATAGGCAAGAAAGATGGTCTCAACACATCAATCTTACTTTTGA 1337
Db      356 uasnAspLeuGlnMetasnAsnSerValasnLeuArgSerIleThrTyLeuValIleas 376
Qy      1338 TGAAGTCGATCGCATGTTGATGAGTTGGTTTGGCCGAAGATGAAGATTAATTTCTTG 1397
Db      376 pgluaIleAspIleMetLeuAspMetGluPheGluProGlnIleArgIleIleLeu----- 394
Qy      1398 CCCGAGAAATCCATCAAGAGAACAGCCCAACCTTATGTCAGTCACTTTCCAGA 1457
Db      395 -----LeuAspValArgProAspArgGlnThrValMetThrSerIleThrProAs 412
Qy      1458 GGAATTCCAAAGTTGGCTGACAGATTTTAAAGTCAATATCTGTTGCTGTTGG 1517
Db      412 pThrValArgGlnLeuAlaLeuSerThrLeuIleuys---AspProMetIleValIleValG 431
Qy      1518 ACAAGTGGGT---GGAGCATGTAGAGATGTTTCAGCAGACCGTTCTCCAAATGGCCAGT 1574
Db      431 yasnLeuIleValIleAlaValasnThrValIleGlnAsnIleIleValIleThrGlyI 451
Qy      1575 CTCGAAAAGAGAAAGCTCGTGAATTCGCGAAACATGAGG---GATGAAGAACTAT 1631
Db      451 sgluIyAspArgAlaLeuThrGlnGlnIleuPheValGluAsnMetSerProAsnAspIleVal 471
Qy      1632 GGTCTTTGTTGAACCTAAGAAAAAGCAGATTTTACTGCACTTTTCTTGTCAAGAAA 1691
Db      471 emetPheValSerGlnIleHisIleAlaAspAspLeuSerSerAspPheAsnIleGlnI 491
Qy      1692 AATATCAACTACAGATATCCATGTCGTCGGAACAGAGACGGAGCAAGCTCTGG 1751
Db      491 yIleSerAlaGlnSerIleuHisIleGlyAsnSerGluGlnSerAspGlnIleValGlnI 511
Qy      1752 AGATTTTCGCTTGAAGAGCCAGCTTCTGTTGCTTACTTCACTGCTCCAGAGGCT 1811
Db      511 uasPheIleuSerGlyAsnIleIleuIleuIleThrAspIleIleValSerArgIle 531
Qy      1812 GGAATTTGAAGATGTCACATGTTTCAATTTGATCTTCTTCAACATGATGATA 1871
Db      531 uasPheIleuAspValIleThrIleValIleTyAsnTyAspPheProArgsnIleAspValI 551
Qy      1872 TGTTCAATCGAATTTGGGGCTACTGCTGCTGTCGGAATACGCGACAGCAATTTCTTTT 1931
Db      551 rValIleAspArgValGlySerIleuAsp-----GlyGlnGlnArgIleuHisIleSerSerI 569
Qy      1932 TGATCTTGAATCGGATACCATTTAGACAGCTCTGTAAGAAAGTATTTGACATGCTCA 1991
Db      569 uIleThrGlnArgAspSerIleSerIleValGlnIleuIleIleIleuAspArgIleAs 589
Qy      1992 ACAGGATGTTCTCT 2004

```

```

Db      589 ngInSerValPro 593
RESULT 10
US-09-815-242-13491
; Sequence 13491, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.01A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13491
; LENGTH: 524
; TYPE: PRN
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13491
US-09-714-865-15 (1-2172) x US-09-815-242-13491 (1-524)
Alignment Scores:
Pred. No.: 3,266-45 Length: 524
Score: 621.00 Matches: 145
Percent Similarity: 57.848 Conservative: 80
Best Local Similarity: 37.288 Mismatches: 144
Query Match: 15.714 Indels: 20
DB: 10 Gaps: 7
Qy      862 CTGACTTTTGAAGACCTAATCTGTGTGACAGACTGAATTAACAACATTCTTAAGCTGT 921.
Db      1 MetIyPheAsnGlnIleuAsnLeuSerAlaAspLeuIleAlaGlnIleGlyIleValG 20
Qy      922 TATACTAACCTAATCTGTCGAAAATACAGATATCCATCATATGTCAGAGACAGAT 991
Db      21 PheValGlnIleSerIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 40
Qy      982 TTGATGCTTGTGCTCAACAGAGGCTGCGAAGACTGCGGCTTTTCTCTACCAATTTG 1041
Db      41 ValIleGlyGlnAlaGlnIleThrGlyIleThrAlaIlePheGlyLeuProThrIle 60
Qy      1042 GCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1101
Db      61 GluIyIleArgThrGlnIleAlaIleThrIle-----Gln 71
Qy      1102 TGATATATGTCAGACCACTGAGAAATGGTCAACACGATTTATTTGGAAGCAGAAAA 1161
Db      72 AlaLeuValIleIleAlaProThrArgGlnIleAlaValGlnSerGlnIleuIleuPheArg 91
Qy      1162 TTTTCTTTGGGACTGTGTAAGAGCTGTTTATATATGAGGGGGAACCCAGCTGGGACAT 1221

```

```

Db      92 PheGlyArgSerLysGlyValLysArgSerValLysGlySerSerIleGluLys 111
Qy      1222 TCATTGCACAAATAGTACCAAGCTGTAAATATATATATGCTACTCTCGAAGACATG 1281
Db      112 GlnIleLysAlaLeuLysSerGlyAlaIleValAlaGlnThrProGlyArgLeuLys 131
Qy      1282 GATATCATAGGCAAAAGAAAGATGCTCTCAACAGATCAAAATCAATGTTTGGATGAA 1341
Db      132 AspieuIleLysArgLysAlaLeuLysLeuGlnAspieuIleGlnThrIleLeuAspieu 151
Qy      1342 GCTGATGCATGTTGATGATGAGTTTGTCCAGAAATGAAAGATTATTTCTGCCCA 1401
Db      152 AlaAspieuIleLeuAsnMetGlyPheLeuGlnAspieuIleGlnAlaIleIleSer----- 169
Qy      1402 GGAATGCCATCAAGAAACAGCCCAACCTTATGTTCACTGCACTTTCCAGAGAA 1461
Db      170 ArgValPro-----GlnAsnArgLysIleThrLeuLeuPheSerAlaThrMetProAspAla 187
Qy      1462 ATTCAAGGTTGGCTGCAGAGATTTTTAAAGTCA---AATATCTGTTTGTGCTGTTGA 1518
Db      188 IleLysArgIleGlyValGlnPheMetLysAlaProGlnIleValLysIleAlaLys 207
Qy      1519 CAAGTGGGTGAGCATGTAGATGTTCAAGCAGACCGTTCTCCAAAGTGGCCAGTTCTCA 1578
Db      208 GluLeu-----ThrThrGluLeuValAspLysIleThrLysValLysGluGln 225
Qy      1579 AAAAGAGAAAAGCTCGTTGAATTTCTCGAAACCTTAGCGATGAAAGACTATGCTTT 1638
Db      226 LysPheAspThrMetThrArgLeuMetAspValAlaGlnProGlnLeuAlaIleValPhe 245
Qy      1639 GTTGAAACTAGAAAAGAGATTTTATCGCAACTTTCTGTTGTCAGAAAATATATCA 1698
Db      246 GlyArgThrLysArgValAspGluLeuThrArgGlyLeuLysIleArgGlyPheArg 265
Qy      1699 ACTACAAAGTATCCATGCTGATCGGAAACAGAGACGGGAGCAAGCTTTGGAGATTTT 1758
Db      266 AlaGluGlyIleHISGlyAspLeuAspGlnAsnLysArgLeuArgValLeuArgAspPhe 285
Qy      1759 CGCTTGGAAAGTCCCAAGTTCTTGTGCTTCACTGATGCTCCAGAGGGCTGGATATT 1818
Db      286 LysAsnGlyAsnLeuAspValLeuValAlaThrAspValAlaAlaAspGlyLeuAspIle 305
Qy      1819 GAAATGTCACAACTGTTATCATTTTGTATGATCTCTTACCATGATGATGATGATGAT 1878
Db      306 SerGlyValThrIleValLysThrAspIleProGlnAspProGlnSerTyrValHis 325
Qy      1879 CGAATGGGCGTACTGCTGTTGGGAAATACTGACAGCAATTCCTTTTGTATCTT 1938
Db      326 ArgIleGlyArgThrGlyArgAlaGlyLysSerGlyIleSerIleThrPheValAlaPro 345
Qy      1939 GAATCGGATAACCTTTAGCACAGCCCTTACTAAAGATTTGACAGATGCTCAACAGAT 1998
Db      346 AsnGlnMetGlyTyrLeu-----GlnIleIleGlnAsnLeuThrLysLysArgMetLys 363
Qy      1999 -----GTTCTGCATGTTGGAGAA 2019
Db      364 GlyLeuLysProAlaSerValGlnGlu 372

RESULT 11
US-09-738-626-4778
; Sequence 4778, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO

```

```

; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4778
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4778

Alignment Scores:
Pred. No.: 6,19e-45 Length: 732
Score: 618.50 Matches: 176
Percent Similarity: 48.078 Conservative: 98
Best Local Similarity: 30.884 Mismatches: 208
Query Match: 15,644 Indels: 88
DB: Gaps: 16

US-09-714-865-15 (1-2172) x US-09-738-626-4778 (1-732)
Qy      580 AATGATGATATCTTTCAGAACAGAACTGCAGTGAAGTGAACGAGTGTATCAAAAGT 639
Db      8 AsnGlyAspValGlnGlnProAsnAsnValIleSerSerGlu----- 21
Qy      640 TTTAATGAAGAACTAATACAGCTCTGGAAGAAAT----- 675
Db      22 ---SerGlnGlnThrProGlnLysAspSerAlaSerAlaAspPheAlaLeuGlnThrPro 40
Qy      676 ---TCTGGAGTCAGACAGACAGAGAGAGAAAGTATGATATCTGACAGACCAAAAGTG 732
Db      41 ThrAsnThrValGlnAspAlaProAlaSerGlnGlySerGlu-----GluIle 56
Qy      733 ACCTACATACCCCTCTCTCAGATGAGATGAGACTCCATCTTTCACATATATCAGACA 792
Db      57 ThrArgVal---AlaAspThrSerGluAspAlaAspSerAlaAspAlaAspAsnAlaSer 75
Qy      793 GGCATA-----AAGTTGCACAAATACGACACTATCTTGTCGAAGTGTCTGACAT 843
Db      76 AsnValIleAsnGlnAsnGlnAspSerSerGlnGlyAlaAsnGlnProSerAsnGlnSer 95
Qy      844 GATGCACACACCAAGTCTGACTTTTGAAGAACTTCTCTGTCGACAGACTGATTAAC 903
Db      96 SerSerThrGlnAlaLysSerGlyPheAspAlaLeuGlyLeuProGlnArgValLeuAsp 115
Qy      904 AACATGCTAAAGCTGTATATCTAAGCTTACTCTGTGCCAAATAATACGATATCTCTATC 963
Db      116 AlaValArgLysValGlyTyrGlnThrProSerProIleGlnAlaGlnThrIleProIle 135
Qy      964 ATACTTCGACGAGACGATTTGATGCTTGTCTCAACACGGGCTGGGAAAGCTGGGCT 1023
Db      136 LeuMetGlnGlyGlnAspValAlaGlyLeuAlaGlnThrGlyThrIleLysThrAlaAla 155
Qy      1024 TTTCTCTCAATTTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 1083
Db      156 PheAlaLeuProIleLeuAlaArgIle-----AspLys 166
Qy      1084 GAGTTCAGAGAACAGATGATATATATGATGATGATGATGATGATGATGATGATGATGAT 1143
Db      167 SerValArgSerProGlnAlaLeuValLeuAlaProThrArgGlnLeuAlaLeuGlnVal 186
Qy      1144 TATTTGGAAGCCGAAATTT---TCTTTGGGACTGTGTA-----AGAGCTGTT 1191
Db      187 -----AlaAspSerPheGlnSerPheAlaAspHisValGlyGlyLeuAsnValLeu 203

```



```

QY 1192 GTTATATATGGGGAACCCAGCTGGACATTCAATTGACAAATAGTACAGGCTGTAAAT 1251
    |||||
DB 204 ProileTyrGlyGlyGlnAlaTyrGlyIleGlnLeuSerGlyLeuArgGlyAlaHis 123
QY 1252 ATATATGCTCTCTCTGGAAGACGTGATATCATAGGCAAGAAAGATTTGGCTC 1311
    |||||
DB 224 LleValValGlyThrProGlyArgIleIleAspHisLeuGlnIleuSerLeuAspIle 243
QY 1312 AACACATCAATATCTAGTTGTTGGATGAAAGTATGATGCTGATGTTGGATGTTGGT 1371
    |||||
DB 244 SerGlyLeuArgPheLeuValIleLeuAspGlnAlaAspGlnLeuLeuSerGlyPheGln 263
QY 1372 CCAGAAATGAAGATTAATTTCTGCCAGAAATCCATCAAGAACGACCCCAACC 1431
    |||||
DB 264 GluAspValGlnArgIleLeuGln-----AspThrProAspGlnIleuVal 279
QY 1432 CTATATGCTGCTGCACTTTTCCAGAGAAATTCAAAGTTGGCTGACAGTTTAAAG 1491
    |||||
DB 280 AlaLeuPheSerIleThrMetProAsnGlyIleArgArgLeuSerGlyIleu--- 298
QY 1492 TCAAAATATCTGTTGCTGTTGGACAAAGTGGTGGACATGATGATGTTCCAGAC 1551
    |||||
DB 299 AsnAsnProAlaGlnIleThrValIleSerGlnThrArgThrAsnThrAsnIleThrGln 318
QY 1552 ACCGTTCTCCAGTTGGCCAGTTCTCAAAAAGAAAAAGCTCGTGAATTCGCAAAC 1611
    |||||
DB 319 ArgPheLeuAsnValAlaHisArgAsnIleuSerGlyLeuThrArgIleLeuVal 328
QY 1612 ATAGGGGAGAAAGAACTATGCTTTGTTGAACATCAAGAAAGAAAGCAAGTTTACTGCA 1671
    |||||
DB 339 ThrGlnPheGlnIleMetIleMetPheValArgThrIleHisGlnIleuThrGlnIleuAla 358
QY 1672 ACTTTCTTGTGCAAGAAAAAATCACTACACATGATGCTGATGCTGGAGAACAGAGA 1731
    |||||
DB 359 GlnIleLeuArgIleArgIlePheSerAlaAlaIleAsnIleuSerIleAlaGlnAla 378
QY 1732 GAGCGGAGACAGCTCTGGAGATTTTCGCTTGGAAAGTCCCACTTCTTGTGTACT 1791
    |||||
DB 379 GlnArgGlnArgThrValAspGlnLeuIleuAspGlyArgLeuAspIleLeuValAlaThr 398
QY 1792 TCGATGCTGCCAGAGGCTGATATTTGAAGATGCAACATGTTCAATTTGATCTT 1851
    |||||
DB 399 AspValAlaIleArgIleuAspValGlnArgIleSerHisValLeuAsnPheAspIle 418
QY 1852 CCTCTACCATGATGATGATGTCATGCAATGGGCGTCTGCTGCTGGGATGATCT 1911
    |||||
DB 419 ProAsnAspThrGlnSerIleValHisArgIleGlyArgThrGlnArgAlaGlyArgThr 438
QY 1912 GGCAGAGCAATTCCTT----- 1929
    |||||
DB 439 GlyGlnAlaIleLeuPheValThrProArgGlnArgArgMetLeuArgSerIleGlnArg 458
QY 1930 -----TTGATCTTGAATCGATAC 1950
    |||||
DB 459 AlaThrAsnAlaProLeuHisGlnMetGlnLeuProThrValAspGlnValAsnAspPhe 478
QY 1951 CAT-----TTAGCAGAGCTCTAGTAAAGTATGACAGATGCTCAACAGATGTT 2001
    |||||
DB 479 ArgIleValIlePheAlaAspSerIleThrIleSerIleuGlnIleuSerGlnMetAspLeu 498
QY 2002 CCTGATGCTGGAGAAATTCCTTATGATACATTCCTGCTGCTCACTGCTAGTACA 2061
    |||||
DB 499 -----PheArgThrLeuValIleuSer----- 507
QY 2062 AGAGGAAAGCTGTTGCTGATGATACAGAAAGGCGAAGACCTTTGAACACAGCT 2121
    |||||
DB 508 GlnAlaAsnAspValProLeuGlnAspIleAlaAlaLeuAlaThrGln-AlaGlnSe 527
QY 2122 GGGT-----TTCTCTTCCAGAGCTCCA 2146
    |||||
DB 527 rGlyAspPheLeuLeuIleuSerGlnLeuPro 536

```

```

US-09-815-242-11308
; Sequence 11308, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Traxick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11308
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11308

Alignment Scores:
Pred. No.: 1,17e-44 Length: 492
Score: 614.50 Matches: 140
Percent Similarity: 56.62% Conservative: 78
Best Local Similarity: 36.36% Mismatches: 142
Query Match: 15.54% Indels: 25
DB: Gaps: 6

US-09-714-865-15 (1-2172) x US-09-815-242-11308 (1-492)
QY 865 ACTTTGAAGAGCTAATCTCTGTGACAGACGTAATTAACACATTGCTAAAGCTGTAT 924
    |||||
DB 21 SerPheAsnAspLeuGlyLeuIleuSerValIleuIleuSerValIleuIleuValIleuVal 40
QY 925 ACTAAGCTTACTCTGCTGCAAAATACTATCTTATCATCTTGCAGACGAGATTGG 984
    |||||
DB 41 ThrSerProSerProIleGlnIleuIleuValIleuValIleuValIleuValIleuValIleuVal 60
QY 985 ATGCTGTGCTCAACAGAGGCTCTGGAAAGACTGGGCGCTTTCTCTTACCATTTGGCT 1044
    |||||
DB 61 IleAlaGlnAlaGlnThrIleGlyThrIleGlyThrIleAlaIlePheAlaLeuProIleIle--- 79
QY 1045 CATATGATCATGATGAAATACTGCGACAGCTTTTAAAGATGTCAGAGAACAGAGTGT 1104
    |||||
DB 80 -----AsnAsnLeuIleuAsnAsnHisThrIleGlnAla 90
QY 1105 ATTATTGACCAACCACTGAGAAATGCTCAACAGATTTATTGGAAACCAAGAAATTT 1164
    |||||
DB 91 LeuValIleThrProThrArgGlnLeuAlaMetGlnIleSerAspGlnIlePheIleu 110
QY 1165 TCTTTGGACCTTGTGTAGACCTGTGTATATATGAGGGAACCAAGCTGGAGACTTCA 1224
    |||||
DB 111 GlyLeuHisThrArgThrIleThrIleValCysValTyrGlyGlnIleuSerValIleuSerValIleuSerVal 130

```

```

QY 1225 ATTCGCAAAATAGTACAGGCTGTATATATATATGCTACTCTCGGAGAGCTGATGAT 1284
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 131 CysGluPheIleIysIysAsnProGlnValMetIleIaIthrProGlyAArgLeuLeuAap 150
QY 1285 ATCATAGGCAAAAGAAAAGATTGGTCTTCAAAACAGATCAATATCTAGTTTGGATGAAGCT 1344
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 HisLeuIysAsnGluArgIleHisIlysPheValProIysValValIleLeuAapGluSer 170
QY 1345 GATCGCATGTGGATGTATGGTTTGTGTCAGAAATGAAGAATTAATTTCTGCCACAGA 1404
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 171 AspGluMetLeuSerMetGlyPheLeuAspAspIleGluGluIlePheAspTyr----- 188
QY 1405 ATGCGATCAAGAACGACGCCCAACCCCTATGTCTAGTCCAACTTTTCCAGAGGAAT 1464
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 LeuProSer-----GluAlaGlnIleLeuLeuPheSerIaIaIthrMetProGluProIle 206
QY 1465 CAAAGCTGGCTGCAGACTTTTAAAGTCAAAATATATGTTTGTCTGTGTCAGCAAGTG 1524
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 207 LysArgLeuIaAspIlysIleLeuGlu---AsnProIleIysIleHisIleAlaProSer 225
QY 1525 GGTGAGCATGTAGAGATGTTCAGACAGCCGTTCTCCAGTTGGCCAGTTCGAAAAAGA 1584
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 226 AsnIleThrAsnIthrAspIleIthrGlnArgPheTyrValIleAsnGlnHisGluArgAla 245
QY 1585 GAAAGCTCGTGAATTTCTGCGAAACATAGAGGAGTGAAGAACTATGCTCTTGTGTGAA 1644
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 246 GluAlaIleMetArgLeuLeuAapThrGlnAlaProIysIysSerIleValPheIthrArg 265
QY 1645 ACTAGAGAAAAGCAGATTTTACTGCAACTTTTCTTGTCAAGAAAAATATCACTACA 1704
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 ThrIysIysGluIaAspArgIleuHisIleGlnPheLeuIaSerIysAsnTyrIlysSerThr 285
QY 1705 AGTATCGATGTGATCGGAAACAGAGAGCGGAGAGCAAGCTCTGAGATTTTTCGTTT 1764
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 AlaLeuHisGlyAspMetAspArgIaArgAspArgSerSerIleMetAlaPheIlys 305
QY 1765 GGAAGTGGCCAGTTCTTGTGTACTTCTAGTAGCTCCAGAGGCTGATATTTGAAAT 1824
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 306 AsnAspIaAspValIleValAlaIaIthrAspValAlaSerArgGlyLeuAapIleSerGly 325
QY 1825 GTGCAACATGTTATCAATTTGATCTTCTTCAATCAATGAAATGTTTCATCGAAT 1884
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 326 ValSerHisValIlePheAsnTyrHisLeuProIysAsnIthrGluSerTyrIleHisArgIle 345
QY 1885 GGGGCTGCTGCTGCTGGGAATACAGACAGCAAGCAATTTCTTTTGAATCTTGAATCG 1944
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 346 GlyArgThrGlyArgAlaGlyIysIysGlyMetAlaIleThr----- 359
QY 1945 GATTAACCATTTAGCAGACGCTCTA---GTAAAGTATTTGACAGATGCTCAACAGATGT 2001
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 360 -----LeuValIthrProIeGluItyrIysIleuLeuAapGluMetGlnIysGluIle 376
QY 2002 CCTGCATGTTTGAA 2016
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 377 AspSerGluIleGlu 381
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

/ TYPE: PR1
/ ORGANISM: Helicobacter sp.
US-09-732-091-2

Alignment Scores:
Pred. No.: 1,17e-44 Length: 492
Score: 614.50 Matches: 140
Percent Similarity: 56.62% Conservative: 78
Best Local Similarity: 36.36% Mismatches: 142
Query Match: 15,54% Indels: 25
DB: 10 Gaps: 6

US-09-714-865-15 (1-2172) x US-09-732-091-2 (1-492)
QY 865 ACTTTGGAAGAGCTAATCTGTCTGACAGACTGAATTAACAACTTGTAAAGCTGTTAT 924
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 SerPheAsnAspLeuGlyLeuIysGluSerValIleuIysSerValItyrGluAlaGlyPhe 40
QY 925 ACTTAAGCTTACTCTGTGCGAAATAATACATATCTTATCATCTTGCAGAGCAGATTGG 984
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 41 ThrSerProSerProIleGlnIuIysAlaIleProAlaValIleuGlnIuIysArgPheVal 60
QY 985 ATGCTGTGCTCAAAACAGAGGCTCTGGAAAGACTGGGCGCTTTCTCTACCAATTTTGGCT 1044
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 IleAlaGlnAlaGlnIthrGlyItyrIthrAlaIaPheAlaLeuProIleIle--- 79
QY 1045 CATATGATCATGATGAATTAATCTGCGACGCTTTTAAAGCTTGCAGAGACCGAGCTGT 1104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 -----AsnAsnLeuIysAsnAsnHisIthrIleGluAla 90
QY 1105 ATTATGTGACACCACTCGAGAAATTTGTCAACAGATTTATTTGAAAGCCAGAAATTT 1164
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 91 LeuValIleIthrProIthrArgIleuAlaMetGlnIleSerAspGluIlePheIysLeu 110
QY 1165 TCTTTTGGACTTGTGTAAAGACTGTGTTATATATGTTGGGGGAACCCAGCTGGACATTGA 1224
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 111 GlyIysHisIthrArgIthrIysThrValCysValItyrGlyGlnIserValIysIysGln 130
QY 1225 ATTCGCAAAATGTCAAGGCTGTATATATATATGTTAGTCTCTGGAAGACTGATGAT 1284
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 131 CysGluPheIleIysIysAsnProGlnValMetIleIaIaIthrProGlyArgLeuLeuAap 150
QY 1285 ATCATAGGCAAAAGAAAAGATTGGTCTTCAAAACAGATCAATATCTAGTTTGGATGAAGCT 1344
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 HisLeuIysAsnGluArgIleHisIlysPheValProIysValValIleLeuAapGluSer 170
QY 1345 GATCGCATGTGGATGTATGGTTTGTGTCAGAAATGAAGAATTAATTTCTGCCACAGA 1404
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 171 AspGluMetLeuAspMetGlyPheLeuAapAspIleGluGluIlePheAspTyr----- 188
QY 1405 ATGCGATCAAGAACGACGCCCAACCCCTATGTCTAGTCCAACTTTTCCAGAGGAAT 1464
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 LeuProSer-----GluAlaGlnIleLeuLeuPheSerIaIaIthrMetProGluProIle 206
QY 1465 CAAAGCTGGCTGCAGACTTTTAAAGTCAAAATATATGTTTGTCTGTGTCAGCAAGTG 1524
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 207 LysArgLeuIaAspIlysIleLeuGlu---AsnProIleIysIleHisIleAlaProSer 225
QY 1525 GGTGAGCATGTAGAGATGTTCAGACAGCCGTTCTCAAGTTGGCCAGTTCGAAAAAGA 1584
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 226 AsnIleThrAsnIthrAspIleIthrGlnArgPheTyrValIleAsnGlnHisGluArgAla 245
QY 1585 GAAAGCTCGTGAATTTCTGCGAAACATAGAGGAGTGAAGAACTATGCTCTTGTGTGAA 1644
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 246 GluAlaIleMetArgLeuLeuAapThrGlnAlaProIysIysSerIleValPheIthrArg 265
QY 1645 ACTTAAGAAAAGCAGATTTTACTGCAACTTTTCTTGTCAAGAAAAATATCACTACA 1704
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 266 ThrIysIysGluIaAspArgIleuHisIleGlnPheLeuIaSerIysAsnTyrIlysSerThr 285
QY 1705 AGTATCGATGTGATCGGAAACAGAGAGCGGAGCAAGCTCTTGCAGATTTTTCGCTTT 1764
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 AlaLeuHisIleGlyAspMetAspArgIaArgAspArgSerSerIleMetAlaPheIlys 305

```

```

QY 1765 GGAAGTCCAGTCTTGTCTACTCTAGTACCTGCAGAGGCTGATATTGAAAT 1824
Db 306 AsnAspAlaAspValLeuValAlaThrAspValAlaSerArgLysLeuAspLysSerGly 325
QY 1825 GTGCAACATGTTATCATTTTGTGATCTTCTTCCATGATGATATGTTGATCAAT 1884
Db 326 ValSerHisValPheAsnIleuProLeuAsnThrGluSerTyrIleHisArgIle 345
QY 1885 GGGCGTACTGCTTGTGGGATATGTCAGAGCAATTTCTTTTGTGATCTTGAATCG 1944
Db 346 GlyArgThrGlyArgAlaGlyLysGlyMetAlaIleThr----- 359
QY 1945 GATTAACATTATGACAGACCTCTA---GTAAAGTATTGACAGATGCTCAACAGATGTT 2001
Db 360 -----LeuValThrProLeuGluTyrLysGluLeuLeuArgMetGlnLysGluIle 376
QY 2002 CTTGCATGTTGGAA 2016
Db 377 AspSerGluIleGlu 381

RESULT 14
US-09-732-091-42
; Sequence 42, Application US/09732091
; Patent No. US20020107368A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Jing-Hui
; APPLICANT: Walker, Richard I.
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses
; FILE REFERENCE: 7969-088
; CURRENT APPLICATION NUMBER: US/09/732,091
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ. ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 42
; LENGTH: 504
; TYPE: PR1
; ORGANISM: Helicobacter sp.
US-09-732-091-42

Alignment Scores:
Pred. No.: 1,18e-44 Length: 504
Score: 614.50 Matches: 140
Percent Similarity: 56.62% Conservative: 78
Best Local Similarity: 36.36% Mismatches: 142
Query Match: 15.54% Indels: 25
Gaps: 6

US-09-714-865-15 (1-2172) x US-09-732-091-42 (1-504)
QY 865 ACTTTGAAGAAGCTATCTCTGTCAGACACTGATATACAAACATTTGCTAAAGCTGTTAT 924
Db 33 SerPheAsnSplLeuGlyLeuLysGlnSerValLeuLysSerValTyrGlnAlaGlyPhe 52
QY 925 ACTAAGCTTACTCTGTCGAAAAATACAGATTTCTTCACTACTGTCAGAGCGAGATTTG 984
Db 53 ThrSerProSerProLysGlnGlnLysLysAlaLeuProAlaValLeuGlnGlyArgAspVal 72
QY 985 ATGGCTTGTGCTCAACAGAGGCTGGAAGACTGCGGCTTTCTCTTCAACAAATTTTGCT 1044
Db 73 IleAlaGlnAlaGlnThrGlyThrGlyLysThrAlaAlaPheAlaLeuProIleLeu--- 91
QY 1045 CATATGATGATGATGATGATATCTCCAGCTGTTTAAAGATTGTCAGAAACAGAGTGT 1104
Db 92 -----AsnAsnLeuLysAsnAsnHisThrIleGluAla 102
QY 1105 ATATATGAGACCACTGAGATTTGTCACCAAGATTTATTTGAGAGCAAGAAATTT 1164
Db 103 LeuValIleThrProThrArgGlnLeuAlaMetGlnIleSerArgGlnIlePheLysLeu 122
QY 1165 TCTTTGGAGCTGTGTAGAGCTGTGTATATATATATGAGGAGAACCCAGCTGGGACATTCA 1224

```

```

Db 123 GlyLysHisThrArgThrLysThrValCysValTyrGlyGlyGlnSerValLysLysGln 142
QY 1225 ATTGCACAAATTTATACAGAGCTGTAATATATATATGCTCTCTCTCGAAGACATGATGAT 1284
Db 143 CysGluPheIleLysLysAsnProGlnValMetIleAlaThrProGlyArgLeuLeuAsp 162
QY 1285 ATATATGAGCAAGAAAGATTTGCTTCAAAACATCAATATCTAGTTTGTGATGAAGCT 1344
Db 163 HisLeuLysAsnGlnArgIleHisLysPheValProLysValValValLeuAspGlnSer 182
QY 1345 GATCGCATGTTGATATGATGTTGTTTGTGTCAGAAATGAAGATTAATTTCTGCCCAAGA 1404
Db 183 AspGluMetLeuAspMetGlyPheLeuAspArgIleGluGlnIlePheAspTyr----- 200
QY 1405 ATGCCATCAAGAGAACAGCCCAACCCCTATGCTACGTCACATTTTCCAGAGGAATTT 1464
Db 201 LeuProSer-----GlnAlaGlnIleLeuLeuPheSerValAlaThrMetProGluProIle 218
QY 1465 CAAGAGTTGCTGTCAGAGTTTAAAGTCAAAATTAATCTGTTGCTGTTGGAACAAGTG 1524
Db 219 LysArgLeuAlaAspLysIleLeuGlu---AsnProLysLysIleHisIleAlaProSer 237
QY 1525 GGTGAGCATGTAGAGATGTTGACAGACCGCTTCCAACTGTCAGATTTCCAAAAAGA 1584
Db 238 AsnIleThrAsnThrAspIleThrGlnArgPheTyrValIleAsnGluHisGluArgAla 257
QY 1585 GAAAGCTCGTTGAATTTCTGCGCAAAACATAGGGGATGAAAGACTATGCTTTGTTGAA 1644
Db 258 GluAlaIleMetArgLeuLeuAspThrGlnAlaProLysLysSerIleValPheThrArg 277
QY 1645 ACTAAGAAAAACAGATTTTACTGCAACTTTTCTTGTGCAAGAAAAATATACACTACA 1704
Db 278 ThrLysLysGlnLysAlaAspGlnLeuHisGlnPheLeuAlaSerLysAsnTyrLysSerThr 297
QY 1705 AGTATCCATGTGATCGGACAGACAGAGACCGGAGCAACCTTGTGAGATTTTCCGTTT 1764
Db 298 AlaLeuHisGlyAspMetAspGlnArgAspArgSerSerIleMetAlaPheLysLys 317
QY 1765 GGAAGTCCCACTTCTTGTGCTACTTCACTGACGTGCGCAGAGGCTGATATTGAAAT 1824
Db 318 AsnAspAlaAspValLeuValAlaThrAspValAlaSerArgLysLeuAspLysSerGly 337
QY 1825 GTCAACATGTTATCATTTTGTGATCTTCTTCACTGATGATGATATGTTGATCGAAT 1884
Db 338 ValSerHisValPheAsnTyrHisLeuProLeuAsnThrGluSerTyrIleHisArgIle 357
QY 1885 GGGCGTACTGCTGTTGTGGAATGACTGCGACAGCAATTTCTTTTGTGATCTTGAATCG 1944
Db 358 GlyArgThrGlyArgAlaGlyLysLysGlyMetAlaIleThr----- 371
QY 1945 GATTAACATTATACAGACCTCTA---GTAAAGTATTGACAGATGCTCAACAGATGTT 2001
Db 372 -----LeuValThrProLeuGluTyrLysGluLeuLeuArgMetGlnLysGluIle 388
QY 2002 CTTGCATGTTGGAA 2016
Db 389 AspSerGluIleGlu 393

RESULT 15
US-09-963-790A-2
; Sequence 2, Application US/09963790A
; Patent No. US20020115161A1
; GENERAL INFORMATION:
; APPLICANT: FARMWICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE DEAD GENE
; FILE REFERENCE: 032301 WD 230
; CURRENT APPLICATION NUMBER: US/09/963,790A
; CURRENT FILING DATE: 2001-09-27
; NUMBER OF SEQ. ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 624

```

```

TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-963-790A-2

Alignment Scores:
Pred. No.: 1,74e-43 Length: 624
Score: 601.50 Matches: 153
Percent Similarity: 51.67% Conservative: 79
Best Local Similarity: 34.08% Mismatches: 153
Query Match: 15.21% Indels: 64
DB: 10 Gaps: 11

US-09-714-865-15 (1-2172) x US-09-963-790A-2 (1-624)

QY 907 ATTGCTAAGCTGGTTACTAGCTTACTCTCTGGCAAAATACAGATTCTTATCATATA 966
DB 9 ValArgLysValAlGlyTyrGluThrProSerProIleGlnAlaGlnThrIleProIleLeu 28
QY 967 CTTCGACGACGAGATTTGATGCTGTGCTCAACAGGCTCGGAAAGACTGCGGCTTTT 1026
DB 29 MetGluGlyGlnAspValValGlyLeuAlaGlnThrGlyThrGlyLysThrAlaAlaPhe 48
QY 1027 CTCCTACCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
DB 49 AlaLeuProIleLeuAlaArgIle-----AspLysSer 59
QY 1087 TTGCAGACGACGAGATTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1146
DB 60 ValArgSerProGlnAlaLeuValLeuAlaProThrArgGlnAlaLeuGlnVal--- 78
QY 1147 TTGGAAGCCAGAAATTT---TCCTTTGGGACTGTGTGA-----AGAGCTGTGTT 1194
DB 79 -----AlaAspSerPheGlnSerPheAlaAspHisValGlyGlyLeuAsnValLeuPro 96
QY 1195 ATATATGGGGGACCCAGCTGGACATTCATTCGACAAATGATGACAGGCTGTAATTTA 1234
DB 97 ILeTyrGlyGlyGlnAlaTyrGlyIleGlnLeuSerGlyLeuArgArgGlyAlaHisIle 116
QY 1255 TTATGTGCTACTCTCGTGAAGACTGATGATGATGATGATGATGATGATGATGATGATGAT 1314
DB 117 ValValGlyThrProGlyArgIleIleAspHisLeuGlnLysGlySerLeuAspLysSer 136
QY 1315 CAGATCAAAATCTTATGTTTGGATGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1374
DB 137 GlyLeuArgPheLeuValLeuAspGlnAlaAspGlnMetLeuAsnMetGlyPheGlnGlu 156
QY 1375 GAAATGAAGAAATTAATTTCTTGGCCAGAAATGCCATCAAAAGAACAGCCCAACCTTT 1434
DB 157 AspValGluArgIleLeuGlu-----AspThrProAspGlnLysGlnValAla 172
QY 1435 ATGTTCACTGCAACTTTCCAGAGAAATTCAAAGGTTGGCTGCGAGAGTTTAAAGTCA 1494
DB 173 LeuPheSerAlaThrMetCysProAsnGlyIleArgAlaGlySerLysGlnTyrLeu--Asn 191
QY 1495 AATTATCTGTTGTTGCTGTTGCTGACACAGCTGGTGAGCATGTAGAGATGTTCAAGACACC 1554
DB 192 AsnProAlaGluIleThrValLysSerGlnThrArgThrAsnThrAsnIleThrGlnArg 211
QY 1555 GTTCTCCAAAGTTGGCCAGCTTCAAAAAGAGAAAAGCTGTTGAAATTTCTCGAAAACATA 1614
DB 212 PheLeuAsnValAlaHisArgAsnLysMetAspAlaLeuThrArgIleLeuGlnValThr 231
QY 1615 GGGGATGAAGAACTATGCTGTTGTTGAACTAGAAAAGAAAGCAATTTTACTGCAACT 1674
DB 232 GluPheGlnAlaMetIleMetCysPheValArgThrLysHisGluThrGluGluValAlaGlu 251
QY 1675 TTTCTTTGCAAAAATAATATCAACTACAGTATCCATGCTGATCGGGAACAGAGAGAG 1734
DB 252 LysLeuArgAlaArgGlyPheSerAlaAlaHisAlaIleAsnGlyAspIleAlaGlnAlaGln 271
QY 1735 CCGGAGCAAGCTCTTGGAGATTTTGGCTTTGGAAAGTGCCCAAGTTCTTGTGCTACTTCA 1794
DB 272 ArgGluArgThrValAspGlnLeuLysAspGlyArgLeuAspIleLeuValAlaThrAsp 291

```

```

QY 1795 GTAGCTGCCAGAGGCTGATATTTGAAATGTCACATGTTATCAATTTGATCTTCC 1854
DB 292 ValAlaAlaArgGlyLeuAspValGluArgIleSerHisValLeuAsnPheAspIlePro 311
QY 1855 TCTACCATGATGAATATGTTTCATCGAATTTGGCGTACTGCTGTTGGGAATTAAGTGC 1914
DB 312 AsnAspThrGluSerTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGly 331
QY 1915 AGAGCAATTTCCCTT----- 1929
DB 332 GluAlaIleLeuPheValThrProArgGluArgMetLeuArgSerIleGluArgAla 351
QY 1930 -----TTGATCTTGAATCGATTAACAT 1953
DB 352 ThrAsnAlaProLeuHisGluMetGluLeuProThrValAspGlnValAsnAspPheArg 371
QY 1954 -----TTACACAGCCTTATGTAAGTAATTTGACAGATGCTCAACAGATGTTCT 2004
DB 372 LysValLysPheAlaAspSerIleThrLysSerLeuGlnAspLysGlnMetAspLeu--- 390
QY 2005 GCATGTTGGAAGAAATGCTCTTATGATCATATTCCTGCTTCAGTGTGATACAGA 2064
DB 391 -----PheArgThrLeuValLysGluTyrSer-----Gln 400
QY 2065 GGAACGCTGTTGCATCATGTTGATACACAGAAAGGCAAGCACTTTGAACAGAGCTGGG 2124
DB 401 AlaAsnAspValProLeuGluAspIleAlaAlaLeuAlaThrGln--AlaGlnSerG 420
QY 2125 T---TTTCTTCTTACAGAGCTCCCA 2146
DB 420 YAspPheLeuLeuLysGluLeuPro 428

```

Search completed: June 10, 2003, 17:02:55  
Job time : 186.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM nucleic - protein search, using frame\_plus\_n2p model

Run on: June 10, 2003, 16:35:50 ; Search time 24 Seconds  
(without alignments)  
5325.549 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 3954  
Sequence: 1 atgggggagatgaagatcgga.....tagatgatgagatcggaat 2172

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 6.0	Ygapext 7.0
Delop 6.0	Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xid  
-O=/cgn2\_1/USPRO.spool/US09714865/rnat.05062003.111759.26063/app.query.fasta\_1.2311  
-DB=Issued\_Patents.AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09714865 @CGN 1.16 @rnat.05062003.111759.26063 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEJURY -NEG SCORES=0 -WAIT -DESBLOCK=100 -LONGLOG  
-DEF TIMEOUT=120 -MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents.AA.\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1215	30.7	662	3	US-09-058-489-15
2	1215	30.7	662	3	US-09-058-489-16
3	1182.5	29.9	660	3	US-09-058-489-18
4	1182.5	29.9	660	3	US-09-058-489-91
5	709.5	17.9	648	4	US-09-183-706-43
6	709.5	17.9	648	4	US-09-567-995-43
7	704	17.8	1261	4	US-09-208-742-4
8	704	17.8	1261	4	US-09-332-295-2
9	704	17.8	1261	4	US-09-709-979-2
10	566.5	14.3	528	4	US-09-134-001C-4262
11	505.5	12.8	411	4	US-09-318-443-8
12	504.5	12.8	415	4	US-09-318-443-2

13	503.5	12.7	411	4	US-09-318-443-6	Sequence 6, Appli
14	488	12.3	403	2	US-08-607-509-4	Sequence 4, Appli
15	488	12.3	403	2	US-08-634-642-4	Sequence 4, Appli
16	488	12.3	403	3	US-08-989-370-4	Sequence 4, Appli
17	485	12.3	403	2	US-08-533-669A-10	Sequence 10, Appli
18	485	12.3	403	2	US-08-607-509-2	Sequence 2, Appli
19	485	12.3	403	2	US-08-454-036-2	Sequence 2, Appli
20	485	12.3	403	2	US-08-634-642-2	Sequence 2, Appli
21	485	12.3	403	3	US-08-989-370-2	Sequence 2, Appli
22	485	12.3	403	4	US-09-183-861-10	Sequence 10, Appli
23	485	12.3	403	4	US-09-022-765-10	Sequence 10, Appli
24	485	12.3	403	5	PCT-US95-05064-2	Sequence 2, Appli
25	473	12.0	455	4	US-09-134-001C-3653	Sequence 3653, Ap
26	469.5	11.9	407	3	US-08-989-370-6	Sequence 6, Appli
27	467.5	11.8	407	3	US-08-989-370-5	Sequence 5, Appli
28	465.5	11.8	491	4	US-09-314-268-3	Sequence 3, Appli
29	463.5	11.7	448	2	US-08-959-749-2	Sequence 2, Appli
30	463.5	11.7	448	4	US-09-351-497-2	Sequence 2, Appli
31	416	10.5	746	4	US-09-149-934-4	Sequence 4, Appli
32	384	9.7	551	4	US-09-503-391-2	Sequence 2, Appli
33	384	9.7	551	4	US-09-503-391-1	Sequence 4, Appli
34	374	9.5	547	3	US-09-176-657-1	Sequence 1, Appli
35	342.5	8.7	859	4	US-09-149-934-1	Sequence 1, Appli
36	334.5	8.5	560	3	US-09-176-657-7	Sequence 7, Appli
37	284.5	7.2	479	4	US-09-149-934-3	Sequence 3, Appli
38	271	6.9	149	4	US-08-679-493A-85	Sequence 85, Appli
39	244.5	6.2	177	4	US-09-318-443-4	Sequence 4, Appli
40	231.5	5.9	370	2	US-08-559-303B-77	Sequence 77, Appli
41	231.5	5.9	370	4	US-09-175-828-77	Sequence 75, Appli
42	231.5	5.9	607	3	US-08-781-891-75	Sequence 75, Appli
43	230.5	5.8	334	4	US-09-463-702A-38	Sequence 38, Appli
44	223	5.6	637	4	US-09-134-001C-5658	Sequence 5658, Ap
45	221.5	5.6	459	4	US-09-134-001C-3178	Sequence 3178, Ap

## ALIGNMENTS

RESULT 1  
US-09-058-489-15 Application US/09058489  
Patent No. 6103866  
GENERAL INFORMATION:  
APPLICANT: Whitehead Institute for Biomedical Research  
APPLICANT: Lahn, Bruce  
TITLE OF INVENTION: Genes in the No. 6103866-Recombining Region of  
FILE REFERENCE: WHI97-08PA  
CURRENT APPLICATION NUMBER: US/09/058,489  
CURRENT FILING DATE: 1998-04-10  
EARLIER APPLICATION NUMBER: 60/041,877  
EARLIER FILING DATE: 1997-04-11  
NUMBER OF SEQ ID NOS: 91  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 15  
LENGTH: 662  
TYPE: PRT  
ORGANISM: Human  
US-09-058-489-15

## Alignment Scores:

Pred. No.: 2,36e-118  
Score: 1215.00  
Percent Similarity: 56.64%  
Best Local Similarity: 43.22%  
Query Match: 30.73%  
Gaps: 20

US-09-714-865-15 (1-2172) X US-09-058-489-15 (1-662)

OY 214 GAGTGTATATAGGAGATATATACATCCACATGGTGGTGGAGTTGGAGTTGGAAGAGT--- 270  
Db 20 AsplevansnserSersrSpsncinsr-----GlyGlyserThralaserLyGlyArg 37

271 -----TTGGAACAGAGCTTTTCAACACAGAGTTGAAGATGAT 315  
Db TyrIleProHISLeuArgAsnArgGluAlaThrArgGly---PheTyrAspLysAsp 56  
Oy 316 AGCTCTGGTTCTCTGAGAGAGCTAGTATGACTGC-----GAA 354  
Db SerSerGlyTyrSerSerSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76  
Oy 355 GATATCCACACAGGAGAGGCTTTTCCAGAGAGCGGCTATCCAGATGAAATAT 414  
Db AsperArgGlyLysSerSerPhePheSerAspArgGlySer----- 90  
Oy 415 TCAGAGCTTACAGGGCCCTACAGAAAGAGGTGAGAGTATTC-----CGAGGTGC 468  
Db 91 -----GlySerArgGlyArgPheAspAspArgGlyArg 101  
Oy 469 CGTGGAGATTTGGTCTAGAAAGTCCAAATATGACTTAGACCGAAGATGATGACG 528  
Db 102 SerAspTyrAspArgGlyGlySer----- 109  
Oy 529 CGCACTGGTGCCTTTTGGTTCTAGAAACCACTATTAACTGACAGGATATGATGAT 588  
Db 110 -----ArgGlyAsp 112  
Oy 589 ACTTCTCAACAGAGAGTGCAGTGCAGT---GAACAGAGGTGTTTACAAAGTTTAAAT 645  
Db 113 -----ArgSerGlyPheGlyLysPheGlyLysArgGlyLys-----Asn 124  
Oy 646 GAAGAGATTAATAACAGCTCTGGAAGAGATTTCTGGAAGTACAGAAAGAGAGAGAA 705  
Db 125 SerArgTyrPyrAspLysSerAspLysAspLysAspLysAspLysAspLysAspLys 138  
Oy 706 AGTAGTACTCAAGACCAAAAGTGAACCTTACATACCCCTCTCCCTGAGAT--- 762  
Db 139 -----ProLeuProPheSerGlyArgLeu 146  
Oy 763 GAGGACCTCATCTTGGACATTATGACAGGAGCAATTCGCAAAATACAGACTTT 822  
Db 147 GluGluGluLeuPheSerGlyLysAsnThrGlyLysAsnThrGlyLysAsnThrGlyLysAspLys 166  
Oy 823 CTGTGAGAGTGTCTGACATGATGACACACAGCAATTCGACTTTGAAAGACTAAT 882  
Db 167 ProValGluAlaThrGlyAsnAsnCySerProHISLysLeuSerPheSerAspValGlu 186  
Oy 883 CTCTGTCAAGACTGAATTAACAACTTGTCTAAAGCTGTTACTTAACTTACTCTGTG 942  
Db 187 MetGlyGluIleIleMetGlyAsnIleGluLeuThrArgTyrThrArgProThrProVal 206  
Oy 943 CAAATAACAGATTCCTATCATCTTCCAGAGAGAGATTTGATGGCTTGGTCAACA 1002  
Db 207 GluLysHISAlaIleProIleLysGluLysAspLeuMetAlaCysHISAlaIleThr 226  
Oy 1003 GGGTCTGGAGAGCTGCGGCTTTTCTCTACCAATTTGGCTCATATGATGATGATGATGA 1062  
Db 227 GlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIleTyrSerAspGly 246  
Oy 1063 -----ATACTGCCAGTGT-----TTTAAAGAGTTGAGGAA 1095  
Db 247 ProGlyGluAlaLeuArgAlaMetCysLysGluArgGlyArgGlyArgGlySerGlnTyr 266  
Oy 1096 CCAGAGTATATTTAGTACCACTGAGAGATTTGCTCAACATATTTATTTGGAAGC 1155  
Db 267 ProIleSerLeuValLeuAlaProThrArgGluLeuAlaValGlnIleTyrGluAla 286  
Oy 1156 AGAAATTTCTTTTGGAGCTGTGTAGAGCTGTGTATATATGAGGGAACCCAGCTG 1215  
Db 287 ArgLysPheSerTyrArgSerArgValArgProCysValTyrGlyGlyAlaAspLys 306  
Oy 1216 GGACATTCATTTGCAAAATAGTACAGGCTTAAATATTTATGTGCTTACTCTGGAAGA 1275  
Db 307 GlyGlnGlnIleArgAspLeuGluArgGlyCysHISLeuLeuValAlaThrProGlyArg 326

Oy 1276 CTGATGATATCATAGCAAGAAAGATTTGCTCTCAACAGATCAAAATCTTAGTTTG 1335  
Db 327 LeuValAspMetCysGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346  
Oy 1336 GATGAAGCTGATGCGATGTTGGATGATGGTTTGGTCCAGAAATGGAAGATTAATTTCT 1395  
Db 347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgAlaValGlu 366  
Oy 1396 TGCCCAAGATTCGCTCAACAAAGCAACGCGCAACCTTATGTTCACTGCACTTTTCCA 1455  
Db 367 GluAspThrMetProProLysGlyValArgHISThrMetCysPheSerAlaThrPhePro 386  
Oy 1456 GAGCAATTCAAAGTGTGGCTGCGAGATTTTAAAGTCAATATGATGTTGTTGCTGT 1515  
Db 387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspLysTyrIlePheLeuAlaVal 405  
Oy 1516 GCAACAGTGGGTGAGAGCATGATGATGTTCCAGACAGACCTTCTCCAGTGGCCACTTC 1575  
Db 406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValAlaTyrValGluGluSer 425  
Oy 1576 TCAAAAGAGAAAGCTCGTGAATTTCTGCAAACTAGG---GATGAAGAGACTATG 1632  
Db 426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrGlyLysAspSerLeuThrLeu 445  
Oy 1633 GTCCTTGTGAACCTAAGAAAGAGAGATTTTACTGCAACTTTTCTTGTCAAGAAAA 1692  
Db 446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHISGluGly 465  
Oy 1693 ATATCAACTCAAGATATCCATGCTGATCGGAAACAGAGAGCGGGAGCAAGCTTTGGA 1752  
Db 466 TyrAlaCysThrSerIleHISGlyAspArgSerGluArgAspArgGluAlaLeuHIS 485  
Oy 1753 GATTTCGCTTGAAGAGTCCCAAGTCTTGTGCTTACTTCAAGTGCAGAGGCTG 1812  
Db 486 GluPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505  
Oy 1813 GATATGAAATGTGCAACATGTTATCAATTTGATCTTCTTCAATGATGAATAT 1872  
Db 506 AspLeuSerAsnValLysHISValIleAsnPheAspLeuProSerAspIleGluGluTyr 525  
Oy 1873 GTTCATGATTTGGGCTTACTGCTGCTGTGGGAATCTGCGACAGCAATTTCTTTT 1932  
Db 526 ValHISArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545  
Oy 1933 GATCTGAATCGGATTAACATTTAGCACAGCTCTTACTTAAGATGACAGATGCTCAA 1992  
Db 546 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuValGluAlaLys 564  
Oy 1993 CAGATGTTCTGATGATGTTGGAAGAAATTCCTTTAGTACATACATTTCTGCTGCTGAGT 2052  
Db 565 GluGluValProSerTyrPheGluAsnMetAlaTyrGlnHISHis-----TyrLys 581  
Oy 2053 GTTAGTCAAGAGAAACGTG-----TTTGACATCAGTTGAT 2088  
Db 582 GlySerSerArgGlyArgSerLysSerSerArgPheSerGlyGlyPheGlyAlaArgAsp 601  
Oy 2089 ACCGAAAGGCAAGACACTTGAACAGAGCTGGGCTTTTCTTCTTCAAGAGCT 2142  
Db 602 TyrArgGln---SerSerGlyAlaSerSerSerSerSerSerSerSerArgAla 618

RESULT 2  
US-09-058-489-16  
; Sequence 16, Application US/09058489  
; Patent No. 6103886  
; GENERAL INFORMATION:  
; APPLICANT: Whitehead Institute for Biomedical Research  
; APPLICANT: Lahn, Bruce  
; APPLICANT: Page, David  
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of  
; TITLE OF INVENTION: the Y Chromosome  
; FILE REFERENCE: WHI97-08PA  
; CURRENT APPLICATION NUMBER: US/09/058, 489  
; CURRENT FILING DATE: 1998-04-10

```

; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 662
; TYPE: PR1
; ORGANISM: Human
; US-09-058-489-16

Alignment Scores:
Pred. No.: 2,36e-118      Length: 662
Score: 1215.00           Matches: 293
Percent Similarity: 56.64%      Conservative: 91
Best Local Similarity: 43.22%    Mismatches: 180
Query Match: 30.73%           Indels: 114
                                Gaps: 20

US-09-714-865-15 (1-2172) x US-09-058-489-16 (1-662)

QY 214 GAGTGTAAAGCAGATTAATACATCCAAATGGGTTTGGAGTTGGAAGAGT--- 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20 AspleuanserSerAspAenGlnSer-----GlyGlySerThrAlaSerLysGlyArg 37
QY 271 -----TTGGAACAGAGGTTTTCACAAACAGCAGGTTTGAAGATGTGAT 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 TyrIleProhIstLeuArgAsnArgGluAlaThrArgGly--PheTyrAspLysAsp 56
QY 316 AGCTCGGTTTTCGGAAGAGCTAGTAATGACGTC-----GAA 354
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 SerSerLysIlePheSerSerLysAspLysAspAlaTyrSerSerPheGlySerArgSer 76
QY 355 GATATATCCAAACGAGCAGAGGAGGTTTTCACAGAGCGGCTATCGAGATGGAATAT 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 AspSerArgGlyLysSerSerPhePheSerAspArgGlySer----- 90
QY 415 TCAGAAAGCTTCAGGGCCATACAGAAAGGTGGAAGAGTATGTTTC-----CGAGTTGC 468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 -----GlySerArgGlyArgPheAspAspArgGlyArg 101
QY 469 CGTGGAGATTTGGTCTAGGAAGTCGCAATATATGACTTAGACCCAGACGAATGATGAC 528
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 SerAspTyrAspGlyIleGlySer----- 109
QY 529 CGCAGCTGGGCTTTTGGTTCTAGAAAGCACAGTATTAAAGTGACAGATATGATGAT 588
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 -----ArgGlyAsp 112
QY 589 ACTTCTCAAGCAGAGAGTGCAGTGAAGT---GAACGAGGTGTTTCAAAAGTTTAAT 645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 -----ArgSerGlyPheGlyLysPheGluArgGlyGly-----Asn 124
QY 646 GAAGAAATTAACAGGCTCTGGAAGAATTTCTGGAAGTCAGACAGAAAGAGAGGAA 705
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 SerArgTyrCysAspLysSerAspGluAspAspTyrSerLys----- 138
QY 706 AGTAGTATACTCAAGACCAAAAGTACCTACATACCCCTCTCCACCTGAGAGAT--- 762
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 -----ProLeuProPheSerGlyArgLeu 146
QY 763 GAGAGCTCATCTTTGACATTTATCAGACAGCATAAATCTCGACAAATACGACATAT 822
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 GlnGlnIleuPheSerGlyGlyAsnThrGlyIleAsnPheGluLysTyrAspAspIle 166
QY 823 CTGTGTGAAGTGTGACATGATGACACACACCAATTTCTGACTTTTGAAGAAGCTAAT 882
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 ProValIleuAlaThrArgLysAsnAsnCysProhIstIleGluSerPheSerAspValGlu 186
QY 883 CTCTGTGACAGACATGAATTAACATGCTTAAGCTGTATATATGACTTACTGCTGTG 942
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 MetCylGluIleIleMetArgAsnIleGluLeuThrArgTyrThrArgProhIleProVal 206
QY 943 CAAATAATACAGTATCTTATCATATCTTGACAGAGAGATTTGATGCTGTCTCAACA 1002

```

```

Db 207 GlnLysIleAlaIleProIleIleLysGlyLysArgAspLeuMetAlaCysAlaGlnThr 226
QY 1003 GGGCTGGGAAAGCTGGGCTTTCTCTACCAATTTTGGCTCATATGATGATGATGATGAT 1062
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 GlySerGlyLysThrAlaAlaPheLeuProIleLeuSerGlnIleTyrSerAspGly 246
QY 1063 -----ATMACTGCCACTGCT-----TTTAAAGAGTTTCAGGAA 1095
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 ProGlyAlaLeuAlaMetLysGluAsnGlyArgTyrGlyArgGlyLysGlnTyr 266
QY 1096 CCAAGCTGATTAATTTGACCAACCACTCGAAGAAATGTCACACAGATTTATTTGAAAGCC 1155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 ProIleSerLeuValIleuAlaProhIleArgGlyLeuAlaValGlnIleTyrGlnGluAla 286
QY 1156 AGAAATTTCTTTGGACTTGTGAAGAGCTGTGTTATATATGAGGGGACCCAGCTG 1215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 ArgLysPheSerTyrArgSerArgValArgProCysValValTyrGlyGlyAlaAspIle 306
QY 1216 GGCATTTCAATTCGACAAATAGTACAGAGCTGTAAATATATATGCTCTCTCGAAGA 1275
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 GlnGlnIleuArgAspLeuGluArgGlyCysHisLeuLeuValAlaThrProGlyArg 326
QY 1276 CTGATGATATCATAGGCAAAAGAAAGATTGTTCTAACAACATCAAAATCTTATGTTTG 1335
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 LeuValAspMetMetGluArgGlyLysIleGlyLeuAspPheCysLysTyrLeuValLeu 346
QY 1336 GATGAGCTGATGCGCATGTTGATGATGATGTTGGTCCGAATGAAAGAGTAAATTTCT 1395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 AspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIleArgGlyIleValGlu 366
QY 1396 TGCCCAAGATGCGCATCAAAAGAAAGAGAGCGCAACCTTATGTTTCAGTCAACTTTTCCA 1455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GlnAspThrMetProProLysGlyValArgHisThrMetMetPheSerAlaThrPhePro 386
QY 1456 GAGAAATTCAAAGTTTGCTGCAAGTTTAAAGTCAAAATTAATCTGTTGTTGCTGTT 1515
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 LysGluIleGlnMetLeuAlaArgAspPheLeu---AspLysTyrIlePheLeuAlaVal 405
QY 1516 GGAAGATGGTGGAGCATGTAGAGATGTTACACAGACCGTTCTCCAGTTGGCCAGTTTC 1575
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 GlyArgValGlySerThrSerGluAsnIleThrGlnLysValValIlePheValGluGluSer 425
QY 1576 TCAGAAAGAGAAAGAGCTCGTTGAATTTCTGCAAAACATAGAG---GATGAAAGAACTATG 1632
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 AspLysArgSerPheLeuLeuAspLeuLeuAsnAlaThrCylLysAspSerLeuThrLeu 445
QY 1633 GTCTTTGTTGAAATGAAGAAAGAAAGAGATTTTACTGCAACTTTTCTTGTCAAGAAAA 1692
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 ValPheValGluThrLysLysGlyAlaAspSerLeuGluAspPheLeuTyrHisGluGly 465
QY 1693 ATATCAACTCAAGTATCCATGATGATCGGAAACAGAGAGAGCGGAGAGCAAGCTTTGGA 1752
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 TyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArgGluAlaLeuHis 485
QY 1753 GATTTTCGCTTGAAGAGTCCAGATCTTGTGTTGCTACTTACAGTAGTGCAGAGGCTG 1812
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 GlnPheArgSerGlyLysSerProIleLeuValAlaThrAlaValAlaAlaArgGlyLeu 505
QY 1813 GATATGAAATATGCAACATGTTATCATATTTGATCTTCTCTACATGATGATAT 1872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 AspIleSerAsnValLysHisValIleAsnPheAspLeuProSerAspIleGluGluTyr 525
QY 1873 GTTCATGATGATGGGGTACTGCTGCTGTGGGAATATACGACAGAGAAATTTCTTTT 1932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 526 ValHisArgGlyIleGlyArgThrGlyArgValGlyAsnLeuGlyLeuAlaThrSerPhePhe 545
QY 1933 GATCTGGAATCGGATTAACATTTAGACAGCGCTCTGTAATAATATATGACAGATGCTCAA 1992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 Asn---GluArgAsnIleAsnIleThrLysAspLeuLeuAspLeuValGluAlaLys 564
QY 1993 CAGAGATTTCTCATAGTTTGAAGAAATTTGCTTATGATACATATCTTGAGCTTCACT 2052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      565 GluGluValProSerTrpLeuGluAsnMetAlaTyrGluHis-----TyrIys 581
Qy      2053 GGTGTCACAGAGGAACGTC-----TTTCATCAGTTGAT 2088
Db      582 GlySerIserArgIysSerIysSerArgPheSerGlyGlyPheGlyAlaArgAsp 601
Qy      2089 ACCGAAGAAGGACAGCACTTTGAACACACAGCTGGTTTCTTTCAGCAGCT 2142
Db      602 TyrArgIn---SerSerGlyAlaSerSerSerPheSerSerSerArgAla 618

RESULT 3
US-09-058-489-18
; Sequence 18, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-089A
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-18

Alignment Scores:
Pred. No.:      6,01e-115      Length:      660
Score:          1182.50      Matches:      278
Percent Similarity: 58.06%      Conservative: 93
Best Local Similarity: 43.51%      Mismatches: 167
Query Match:    29.91%      Indels:      101
DB:             3           Gaps:          18

US-09-714-865-15 (1-2172) x US-09-058-489-18 (1-660)
Qy      301 TTGAGATGCTAGTCTGCTGTTCTGAGAGAGTCTAGTAATGACTGC----- 351
Db      52 PheHisAspIysAspSerSerArgIYTrpSerCySerIysAspIysAspAlaTyrSerSer 71
Qy      352 -----GAAGATTAATCAACACGAAACAGAGGGTTTCCAGAGAGCGGCTATCGA 402
Db      72 PheGlySerArgAspSerArgGlyIysProGlyTyrPheSerGluArgGlySer----- 89
Qy      403 GATGAATATATTTCAGAAAGCTTCAGAGGCCATACAGAAAGGTGGAAGAGTATGTTTC--- 459
Db      90 -----GlySerArgGlyArgPheAsp 96
Qy      460 ---CGAGGTCCGTCGAGAGATTGGTCTAGAAAGTCCAATAATGACTTAGACCCAGAC 516
Db      97 AspArgGly-----ArgSerAspTyrAsp----- 104
Qy      517 GAATGTATGAGCGCACTGTGGCTTTTGGTTCTAGAGACAGCATTAATAGTGCA 576
Db      105 -----GlyIle 106
Qy      577 GGTATGTGATCTTCTCAAGCAGAAAGTGCAGTGAAGT---GAACGAGGTGTTAC 633
Db      107 GlyAsn-----ArgIysArgProGlyPheGlyArgPheGluArgSerGlyHis 122
Qy      634 AAGGTTAATGAAGAGTATATAACAGCTCTGGAAGAAATTTCTGGAAGTCAAGAAC 693
Db      123 SerArgTrpCyAspIys-----SerValGluAspAspTrpSerIys----- 136
Qy      694 GAAAGAGGAAAGTAGTATATCAAGAACCAAAAGTAGACTATATCCCTCTCTCA 753

```

```

Db      137 -----ProIeuProPro 140
Qy      754 CTTGAGAT---GAGACGTCATCTTTGACATTATACAGACAGCAATTAACCTTCAGACAA 810
Db      141 SerGluArgLeuGluGlnGlnIleuPheSerGlyGlyAsnThrGlyIleAsnPheGlyIys 160
Qy      811 TAGACACTATTTCTTGGAGAGTCTGTGACATATGACACACACCAATTCAGACTTT 870
Db      161 TyrAspAspIleProValGluIleThrGlySerAsnCySerProHisIleGluAsnPhe 180
Qy      871 GAAGAAGCTAATCTCTGTACAGACTGATTAACAACTTGCTTAAGCTGTATTAAG 930
Db      181 SerAspIleAspMetGlyGluIleIleMetGlyAsnIleGluIleuThrArgTyrThrArg 200
Qy      931 CTTACTCTGTGAAAATACAGTATTCATATTCATCTTGCAGACAGAGATTGATGAGCT 990
Db      201 ProThrProValGlnHisAlaIleProIleIleIysGlyIysAspAspLeuValAla 220
Qy      991 TGTGCTCAACAGAGCTGTGGAGAACTGCGCTTTTCTCTACCAATTTTGCTCATATG 1050
Db      221 CysAlaGlnThrGlySerGlySerThrAlaAlaPheIleuIeuProIleuSerGlnIle 240
Qy      1051 ATGCATGATGAATTAAGTCCAGTCCGTTTAAAGAGTTGCAGAA----- 1095
Db      241 TyrThrAspGly---ProGlyGluAlaLeuIysAlaValIysGluAsnGlyArgTyrGly 259
Qy      1096 -----CCAGAGTATATATGTAGCACCAACTGAGAAATGTGCAACAG 1140
Db      260 ArgArgIysGlnTyrProIleSerLeuValIleuAlaProThrArgGluLeuAlaValGln 279
Qy      1141 ATTTATTTGAGAACCCAGAAAATTTCTTTGGAGCTTGTTGTAAGCTGTTGTTATATAT 1200
Db      280 IleTyrGluGlnIleAlaArgIysPheSerTyrArgSerArgValArgProCysValAlaTyr 299
Qy      1201 GGGGGAACCCAGCTGGGACATTCATTCGAAATAGTACAAAGCTGTAATATATATGT 1260
Db      300 GlyGlyAlaAspIleGlyGlnGlnIleArgAspLeuGluIysGlyCysHisLeuLeuVal 319
Qy      1261 GCTACTCTGTGAAGACTGATGATATCATATGCGAAAGAAAGATGTCCTCAACAGATC 1320
Db      320 AlaThrProGlyArgLeuValAspMetGluArgGlyIysIleGlyLeuAspPheCys 339
Qy      1321 AAATACTTACTTTGGATGAAGCTGATCGCATGTTGATATGAGTTTGGTCCAGAAATG 1380
Db      340 LysTyrIleuValIleuAspGluIleAspArgMetLeuAspMetGlyPheGluProGlnIle 359
Qy      1381 AAGAAGTTAATTTCTTGGCCAGAAATGCCATCAAGAGACAGGCCAAACCTTATATGTT 1440
Db      360 ArgArgIleValGluGlnAspThrMetProProIysGlyValAlaArgHisThrMetMetPhe 379
Qy      1441 AGTGCAACTTTTCAGAGAAATTCAAAGGTGGCTGCAGAGTTTAAAGTCAAAATAT 1500
Db      380 SerAlaThrPheProIysGluIleGlnMetLeuAlaArgAspPheLeu---AspGluTyr 398
Qy      1501 CTGTGTGTCGTGTGGCAAGTGGGTGAGCATGTAGAGATGTTCCAGACCGTTTCTC 1560
Db      399 IlePheLeuAlaValGlyArgValGlySerThrSerGluAsnIleThrGlnIysValVal 418
Qy      1561 CAAGTGGCCAGTTCTCAAAAAGAGAAAGACCTGTTGAATTTCTGCCAAACATAGGS--- 1617
Db      419 TrpValGluAspLeuAspIysArgSerPheIleuAspIleuGlnIylalThrGlySer 438
Qy      1618 GATGAAGAAGACTATNGCTTGTGTAACATAAGAAAAGACAGATTTTATGCAACTTT 1677
Db      439 AspSerIeuThrIeuValPheValGluThrIysGlyValAspSerLeuGluAspPhe 458
Qy      1678 CTTTGTCAAGAAAAATATCACTACAGATATCATGTGATCGGAAACAGAGAGACCGG 1737
Db      459 LeuTyrHisGluGlyTyrAlaCysThrSerIleHisGlyAspArgSerGlnArgAspArg 478
Qy      1738 GAGCAAGCTCTTGAGAGATTTTCCCTTGGAAAGTCCCAGCTTGTGTCTACTTCAGTA 1797
Db      479 GluGluAlaLeuHisGlnPheArgSerGlyIysSerProIleIeuValAlaThrAlaVal 498

```



```

QY 1798 GCTGCCAGAGGGCTGATATTGAAATGTGCACATGTTATCAATTTGATCTTCTTCT 1857
Db 499 AlalalarglyleuasapiliseserAsnValalrightisallleasnpheaspleuProSer 518
QY 1858 ACCATTGATGAATATGTTTCATCGAATTTGGCGCTACTGCTGCTTTGGGAATATCGGCAGA 1917
Db 519 AsplleugluTyrValHisargileglyYargThrGlyargValalGlyAsnleuglyleu 538
QY 1918 GCATTTCTTTTGTGATCTTGATCGATACATTAACATTTAGACAGCCCTAGTAAAGTA 1977
Db 539 AlathserPhepneAsn---GluLysAsnMetCsnlleThrLysAspLeuLeuaspleu 557
QY 1978 TTGACAGATGCTCAACAGAGATTTCTGCATGTTTGGAGAATATGCTTTAGTACATAC 2037
Db 558 leuValalGluLysGlnGluValProSerThrleuGluAsnMetAlaTyrGluHisHis 577
QY 2038 ATTCTGGCTTCACT---GGTAGTACAAGAGAAACGTGTTGCA-----TCA 2082
Db 578 TyrlyseglyserArgGlyArgserLysSerAsnArgPheSerGlyGlyPheGlyAla 597
QY 2083 GTTGATACCGAAGGGGAGAGACCTTTGAACACAGCTGGTCTTTCTTTCACGA 2139
Db 598 ArgAspTyrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615

RESULT 4
US-09-058-489-91
; Sequence 91, Application US/09058489
; Patent No. 6103886
; GENERAL INFORMATION:
; APPLICANT: Whitehead Institute for Biomedical Research
; APPLICANT: Lahn, Bruce
; APPLICANT: Page, David
; TITLE OR INVENTION: Genes in the No. 6103886-Recombining Region of
; FILE REFERENCE: WHI97-089A
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 91
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Human
US-09-058-489-91

Alignment Scores:
Pred. No.: 6,01e-115 Length: 660
Score: 1182.50 Matches: 278
Percent Similarity: 58.06% Conservative: 93
Best Local Similarity: 43.51% Mismatches: 167
Query Match: 29.91% Indels: 101
DB: 3 Gaps: 18

US-09-714-865-15 (1-2172) x US-09-058-489-91 (1-660)
QY 301 TTGAAGATGTGATAGCTGTGTTCTGAGAGAGTCTAGTATGACTGC----- 351
Db 52 PheHisAspLysAspSerSerGlyTyrSerCyserLysAspLysAspAlaTyrSerSer 71
QY 352 -----GAAGATATCAACAGAGGAGGAGGTTTCCAGAGAGCGGCGCATATGA 402
Db 72 PheGlySerArgAspSerArgGlyLysProGlyTyrPheSerGluArgGlySer----- 89
QY 403 GATGGAATATATTCAGAGCTTCAGAGGCATACAGAAAGATGAGAGGATAGTTTC--- 459
Db 90 -----GlySerArgGlyArgPheAsp 96
QY 460 ---CGAGTTGCCGTGAGAGATTGGCTCAGAGAGTCCAAATATATGACTTAGACCCAGAC 516
Db 97 AspArgGly-----ArgSerAspTyrAsp----- 104

```

```

QY 517 GAATGTATGACGCGCACTGGTGGCTTTTGGTTCTAGAAAGACGATTAAGTGGACA 576
Db 105 -----Gly11e 106
QY 577 GGTATATGTATCTTCTCAAGAGAGATGGCAGTGGAAGT---GAACGAGTGCTTAC 633
Db 107 GlyAsn-----ArgGluArgProGlyPheGlyArgPheGluArgSerGlyHis 122
QY 634 AAAGCTTAATAGAGAAGATATATACAGGCTCTGGAAGAAGATTTGGAAGTACGAAGCA 693
Db 123 SerArgTyrCyAspLys-----SerValGluAspTyrPheSerLys----- 136
QY 694 GAAAGAGAGAAAGTACTGATATCTCAAGACCAAAAGTGAACCTACATACCCCTCTTCA 753
Db 137 -----ProLeuProPro 140
QY 754 CTTAGAGAT---GAGACCTCATCTTTGCACATTAATACAGACGATTAATCTTGACAAA 810
Db 141 SerGluArgleuGlnGlnGluLeuPheSerGlyGlyAsnThrGlyLeAsnPheGluLys 160
QY 811 TAGGACACTATTTCTTGGAAGTGTCTGCACATGATGACACGACGAAATCTGACTTTT 870
Db 161 TyrAspAspIleProValalGluAlaThrGlySerAsnCySProProHisleGluAsnPhe 180
QY 871 GAAGAGCTATATCTCTGCACACACTGAATACACATTTGCTAAAGCTGTTATCTAG 930
Db 181 SerAspIleAspMetGlyGluIleleMetClyAsnleleGluLeuThrArgTyrThrArg 200
QY 931 CTTACTCTGTGCAAAAATACAGATATTCATCATCTTGCAGACGAGATTTGATGGCT 990
Db 201 ProThrProValalGlnLysHisAlaIleProIleleLysGlyLysArgAspLeuValAla 220
QY 991 TGTGCTCAACAGGCTGTGGGAAGCTGGCTTTCTCCACCAATTTGGCTCATATG 1050
Db 221 CysAlaGlnThrGlySerGlyLysThrAlaAlaPheLeuLeuProIleLeuSerGlnIle 240
QY 1051 ATGCATATGGAATAACTGCGACGCTTTTAAAGATGTCAGAAA----- 1095
Db 241 TyrThrAspGly---ProGlyGluAlaLeuLysAlaValLysGluAsnGlyArgTyrGly 259
QY 1096 -----CGAGAGTATATTGTATGACCACTGACCAATGTCAGAAATGTCACCAAG 1140
Db 260 ArgArgLysGlnTyrProIleSerLeuValleuAlaProThrArgGluLeuAlaValGln 279
QY 1141 ATTTATTTGAGAGCCAGAAAATTTCTTTGGACCTGTGAAGCTGTTATATATAT 1200
Db 280 IleTyrGlnGluAlaArgLysPheSerTyrArgSerArgValArgProCyAspValValTyr 299
QY 1201 GGGGAAACCCAGCTGGGACATTCATTCGACAAATAGTACAAGCTGTATATATTATGT 1260
Db 300 GlyValalAspIleGlyGlnGlnIleArgAspLeuGluArgGlyCyHisleuLeuVal 319
QY 1261 GCTACTCTGTGAGACGTGATGATATCATAGCAAGAAAGAAAGATTGCTCAACAGATC 1320
Db 320 AlathrProGlyArgGluValalAspMetCyluArgGlyLysIleGlyLeuAspPheCyS 339
QY 1321 AAATACTTATGTTTGGATGAGCTGATGTCATGTTGATATAGGCTTTGGTCCAGAAATG 1380
Db 340 LysTyrLeuValleuAspGluAlaAspArgMetLeuAspMetGlyPheGluProGlnIle 359
QY 1381 AAGAAGTAATTTCTTGGCCAGATTCGACCAATAGTACAAGAAAGCGCAACCTTATGTTTC 1440
Db 360 ArgArgIleValalGluGlnAspThrMetProProLysGlyValalArgHisThrMetCylhe 379
QY 1441 AGTGCACTTTTTCAGAGAAATTCAGAGTTGGCTGCGAGATTTTAAAGTCAAAATAT 1500
Db 380 SerAlathrPheProLysGluIleGlnMetLeuAlaArgPheLeu---AspLunTyr 398
QY 1501 CTGTTTGTGCTGTTTGGACAAGTGGTGAGACATGTAGAGATGTTTCAGACACCTTCTC 1560
Db 399 IlePheLeuAlaValalGlyArgValalGlySerThrSerGluAsnIleThrGlnLysValal 418

```

QY 1561 CAAGTGGCCAGTCTTCAAAAAAGAGAAAAAGCTCGTTGAATTCGCGAAATAGG--- 1617  
 Db |||||  
 QY 419 TrryAlGluAspLeuAspLysArgSerPheLeuLeuAlaIleuGlyAlaThrGlySer 438  
 Db |||||  
 QY 1618 GATGAAGAAGTATGCTTTTGTGGAACCTAAGAAAAAGACAGATTTCACGAACTTTT 1677  
 Db |||||  
 QY 439 AsperLeuThrLeuValPheValGluThrLysGlyAlaAspSerLeuGluAspPhe 458  
 Db |||||  
 QY 1678 CTTTGTCAAGAAAAATATCACTACATCAAGTATCCATGCTGATCGGAAACAGAGAGCGG 1737  
 Db |||||  
 QY 459 LeuYrHisGluGlyYrAlaCysThrSerLeuHisGlyAspArgSerGlnArgAspArg 478  
 Db |||||  
 QY 1738 GAGCAAGCTCTTGAAGATTTTCGCTTTGGAAGAGTCCAGTCTTGTCTTACTTCACTA 1797  
 Db |||||  
 QY 479 GluGluAlaLeuHisGlnPheArgSerGlyLysSerProIleLeuValAlaIleThrAlaVal 498  
 Db |||||  
 QY 1798 GCTGCCAGAGGGCTGATATTTGAATGCAACATGTTATCAATTTGATCTTCTTCT 1857  
 Db |||||  
 QY 499 AlaAlaArgGlyLeuAspLysSerAsnValArgHisValIleAsnPheAspLeuProSer 518  
 Db |||||  
 QY 1858 ACCATTTGATGAATATGTCATCAATTCGAGTACCTGCTGTTGGGAAATACGACGA 1917  
 Db |||||  
 QY 519 AspIleGluGlyYrValHisArgIleGlyArgThrGlyArgValGlyAsnLeuGlyLeu 538  
 Db |||||  
 QY 1918 GCAATTTCTTTTGTGATCTTGAATCGGATACCATTTAGCAACAGCTCTTAAAGTA 1977  
 Db |||||  
 QY 539 AlaThrSerPhePheAsn---GluLysAsnMetAsnIleThrLysAspLeuLeuAspLeu 557  
 Db |||||  
 QY 1978 TTGACAGATGCTCAACAGAGATGTTCTGATGTTGGAAGAAATTCCTTTAGTACATAC 2037  
 Db |||||  
 QY 558 LeuValGluAlaLysGlnGluValProSerTrpLeuGluAsnMetAlaYrGluHisHis 577  
 Db |||||  
 QY 2038 ATTCCTGGCTTCAGT---GGTAGTACAGAGAAACGCTTTGCA-----TCA 2082  
 Db |||||  
 QY 578 TyrLysGlyGlySerArgGlyArgSerLysSerAsnArgPheSerGlyGlyPheGlyAla 597  
 Db |||||  
 QY 2083 GTTGATACAGAAAGGCGACAGACATTGAAACACAGCTGGGTTTCTTCTTCAAGA 2139  
 Db |||||  
 QY 598 ArgAspYrArgGlnSerSerGlySer---SerSerSerGlyPheGlyAlaSerArg 615  
 Db |||||

## RESULT 5

US-09-183-706-43  
 ; Sequence 43, Application US/09183706  
 ; Patent No. 6245525  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marcelange, Valrie  
 ; APPLICANT: De Smet, Charles  
 ; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
 ; FILE REFERENCE: L0461/7054  
 ; CURRENT APPLICATION NUMBER: US/09/183,706  
 ; CURRENT FILING DATE: 1998-10-30  
 ; EARLIER APPLICATION NUMBER: 09/122,989  
 ; EARLIER FILING DATE: 1999-07-27  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SEQ ID NO 43  
 ; LENGTH: 648  
 ; TYPE: PRT  
 ; ORGANISM: H. sapiens  
 US-09-183-706-43

## Alignment Scores:

Pred. No.: 2,356-65 Length: 648  
 Score: 709.50 Matches: 160  
 Percent Similarity: 60.00% Conserves: 77  
 Best Local Similarity: 40.51% Mismatches: 145  
 Query Match: 17.94% Indels: 13  
 DB: 4 Gaps: 8

US-09-714-865-15 (1-2172) x US-09-183-706-43 (1-648)

QY 850 CCACCGCAATTCGACTTTTGAAGAACTAATCTGT---CAGACACTGAATAACAC 906  
 ||| |||||:||||| ||| :|||

Db 237 ProAsnProThrCysThrPheAspAspAlaPheGlnCysTyrProGluValMetGluAsn 256  
 QY ATTGCTAAAGCTGGTTATTAAGCTTACTCTGTGCAAAAAATACATATTCATCATATA 966  
 Db |||||  
 QY 257 IleLysLysAlaGlyPheGlnLysProThrProIleIleSerGlnAlaIleProIleVal 276  
 Db |||||  
 QY 967 CTTGACAGACAGATTTGATGCTTGTGTCAAACAGGCTGTGGAAAGACTCGCGCTTTT 1026  
 Db |||||  
 QY 277 LeuGlnIleLysLeuIleGlyValAlaGlnThrGlyThrGlyLysThrLeuCysTyr 296  
 Db |||||  
 QY 1027 CTTCTACCAATTTTGGCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1086  
 Db |||||  
 QY 297 LeuMetProGlyPheIleHisLeuVal-----LeuIleProSerLeuLysGlyGln 313  
 Db |||||  
 QY 1087 TTGACAGAACCAAGATATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1146  
 Db |||||  
 QY 314 ArgAsnArgProGlyMetLeuValLeuThrProThrArgGluLeuAlaLeuGlnValGlu 333  
 Db |||||  
 QY 1147 TTGAGACCAAGAAATTTTCTTTTGGACTGTGTAAAGCTGTGTTATATATGCGGGA 1206  
 Db |||||  
 QY 334 GlyGluCysCysLysTyrSerTyr---LysGlyLeuArgSerValCysValTyrGlyGly 352  
 Db |||||  
 QY 1207 ACCAGCTGGACATTTCAATTCGACAAATAGTACAGGCTGTATATATATGCTACT 1266  
 Db |||||  
 QY 353 GlyAsnArgAspGluGlnIleGluGluLeuLysGlyValAlaAspIleIleAlaThr 372  
 Db |||||  
 QY 1267 CTTGAGACCTGATGATGATATCATAGGCAAGAAAGATTTGCTCAACAGATCAATATAC 1326  
 Db |||||  
 QY 373 ProGlyArgLeuAsnAspLeuGlnMetSerAsnPheValAsnLeuLysAsnIleThrTyr 392  
 Db |||||  
 QY 1327 TTAGTTTGGATGAAGCTGATCGCATGTTGATATGCGTTTGGTCCGAAATGAAAGAG 1386  
 Db |||||  
 QY 393 LeuValLeuAspGluAlaAspLysMetLeuAspMetGlyPheGluProGlnIleMetLys 412  
 Db |||||  
 QY 1387 TTAATTTCTTCCCGACGAATGCCATCAAGAAACAGGCCAACCCATTATGTCAGTCA 1446  
 Db |||||  
 QY 413 IleLeu-----LeuAspValArgProAspArgGlnThrValMetThrSerAla 428  
 Db |||||  
 QY 1447 ACTTTTCAGAGAAATTCAAAGGTTGGCTGACAGAGTTTAAAGTCAAAATATCTGTTT 1506  
 Db |||||  
 QY 429 ThrTrpProHisSerValHisArgLeuAlaGlnSerTyrLeuLysGluPro--MetIle 447  
 Db |||||  
 QY 1507 GTTGCTGTGACAGAGTGGT---GGAGCATGTAGAGATGTTGACAGACCGTTCGAA 1563  
 Db |||||  
 QY 448 ValTyrValGlyThrLeuAspLeuValAlaValSerSerValLysGlnAsnIleIleVal 467  
 Db |||||  
 QY 1564 GTTGCCAGTTCCTCAAAAAGAAAGAAAGCTGTTGAATTTCCGAAACATAGGGAT--- 1620  
 Db |||||  
 QY 468 ThrThrGluGluGlnLysTrpSerHisMetGlnThrPheLeuGlnSerMetSerSerThr 487  
 Db |||||  
 QY 1621 GAAAGAACTATGCTTGTGTAACCTAAGAAAAAGCAGATTTTACTGCACTTCTTCTT 1680  
 Db |||||  
 QY 488 AspLysValIleValPheValSerArgLysAlaValAlaAspHisLeuSerSerAspLeu 507  
 Db |||||  
 QY 1681 TGTCAAGAAAAATATCACTACATCAAGTATCCATGCTGATCGGAAACAGAGAGCGGAG 1740  
 Db |||||  
 QY 508 IleLeuGlyAsnIleSerValGlnSerLeuHisGlyAspArgGluGlnArgAspArgGlu 527  
 Db |||||  
 QY 1741 CAGGCTTTGAGACTTTTCGCTTTGCAAGAGTCCGACTTCTGTGCTACTTCACTAGTACT 1800  
 Db |||||  
 QY 528 LysAlaLeuGluLysAsnPheLysThrGlyLysValArgIleLeuIleAlaThrAspLeuAla 547  
 Db |||||  
 QY 1801 GCCAGAGAGGCTGATATTTGAATGTCGAACATGTTTCAATTTGATCTTCTTCAACC 1860  
 Db |||||  
 QY 548 SerArgGlyLeuAspValHisAspValThrHisValLysAsnPheAspPheProArgAsn 567  
 Db |||||  
 QY 1861 ATTGATGAATATGTTCAATTCAGATTTGGGCTGATCGTGTGTTGGGAATACCTGACAGACA 1920  
 Db |||||  
 QY 568 IleGluGluYrValHisArgIleGlyArgThrGlyArgAlaGlyArgValSer 587  
 Db |||||  
 QY 1921 ATTTCCTTTTGTGATCTTGAATCGGATTAACCATTTAGCACAGCTCTTGTAGTAAAGTATG 1980  
 Db |||||  
 QY 588 IleThrThrLeu---ThrArgAsnAspTrpArgValAlaSerGlnLeuIleAsnIleLeu 606  
 Db |||||

```

QY      1981 ACAGATGCTCAACAGATGTTCTCGATGCTGGAAGAAATTGCC 2025
DB      607 GUARGAlaasnGInserileProGluGluLeuValserMetala 621

RESULT 6
US-09-567-995-43
; Sequence 43, Application US/09567995
; Patent No. 6303756
; GENERAL INFORMATION:
; APPLICANT: Martelange, Val,rie
; APPLICANT: De Smet, Charles
; APPLICANT: Boon-Falloux, Thierry
; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: L0461/7054
; CURRENT APPLICATION NUMBER: US/09/567,995
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/183,706
; NUMBER OF SEQ ID NOS: 43
; SEQ ID NO 43
; LENGTH: 648
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-567-995-43

Alignment Scores:
Pred. No.:      2,35e-65      Length:      648
Score:          709.50      Matches:      160
Percent Similarity: 60.00%      Conservative: 77
Best Local Similarity: 40.51%      Mismatches: 145
Query Match:    17.94%      Indels:      13
DB:             4           Gaps:          8

US-09-714-865-15 (1-2172) x US-09-567-995-43 (1-648)

QY      850 CCACCGCAATTTGACTTTTGAAGAACTAATCTCTG--CAGACTGAAATAACAC 906
DB      237 ProAsnProThrCysThrPheAspAlaPheGluCysTyrProGluValMetGluAsn 256

QY      907 ATTGTAAGCTGCTTAATAGCTTACTCCTGCGCAAAAATACGATATTTCTATCATA 966
DB      257 lIeLysLysAlaIlyPheGluInLysProThrProIleGlnserGlnAlaTrrProIleVal 276

QY      967 CTTCGACGACGAGATGATGCTTGTGCTCAACAGGCTGCGAAGACTGCGGCTTT 1026
DB      277 LeuGlnGlyIleAspLeuIleGlyValAlaGlnThrGlyIleThrLeuCysTyr 236

QY      1027 CTCTACCAATTTGGCTCATATGATGATGATGAATAATCCAGTCGTTTAAAG 1086
DB      297 LeuMetProGlyPheIleIleIleVal-----LeuGlnProSerLeuLysGlyGln 313

QY      1087 TTGCAGGAACCGAGGTATATATGTAGACCAACCTCGAATTTGGTCAACCGATTTAT 1146
DB      314 ArgAsnArgProGlyMetLeuValLeuThrProThrArgGluLeuValLeuGlnValGlu 333

QY      1147 TTGGAACGACGAATAATTTCTTTTGGAGCTTGTGTAGAGACTGTTGTTATATAGGGGA 1206
DB      334 GlyGluCysCysLysTyrSerTyr---LysGlyLeuArgSerValCysValTyrGlyGly 352

QY      1207 ACCCAGCTGGGACATTCATTCGACAAATAGTACAAAGCTGTAAATATTATGTGCTACT 1266
DB      353 GlyAsnArgAspGluGlnIleGluGlnLeuLysGlyValAlaSerIleIleIleValThr 372

QY      1267 CCTGGAAGACTAGTATCATATGACGCAAAAGAAATTGCTTCAACGATCAATATAC 1326
DB      373 ProGlyArgLeuAsnAspLeuGlnMetSerAsnPheValAsnLeuLysAsnIleThrTyr 392

QY      1327 TTAGTTTGGATGACCTGATCGCATGTTGATATGAGTTTGGTCAGAAATGAAGAAG 1386
DB      393 LeuValLeuAspGluIleAspLysMetLeuAspMetGlyPheGluIleProGlnIleMetLys 412

QY      1387 TTAATTTCTTGCCCAAGATGCCATCAAAAGAAACAGCGCAACCTTATGTTCAAGTGA 1446

```

```

DB      413 lIeLeu-----LeuAspValArgProAspArgGlnThrValMetThrSerAla 428
QY      1447 ACTTTCCAGAGAAATTCAAAGTTGGCTGCGAGAGCTTTTAAAGTCAATTAATTCGTTT 1506
DB      429 ThrTrpProHisSerValHisArgLeuAlaGlnSerTyrLeuLysGluPro--MetIle 447
QY      1507 GTTCTGTTGGACAGATGGGT--GGAGCATGTAGAGATGTTGACGACGCTTCTCAA 1563
DB      448 ValTyrValGlyIleThrLeuAspLeuValAlaValSerSerValLysGlnAsnIleIleVal 467
QY      1564 GTTGGCCAGTTCTCAAAAAGAGAAAGCTGTTGAAATTTGGCAACATAGGGGAT-- 1620
DB      468 ThrThrGluGluGluLysTyrSerHisMetGlnThrPheLeuGlnSerMetSerThr 487
QY      1621 GAAGAAGACATAGTCTTGTGGAACATAGAAAGAAAGAGATTTTACTGCACTTTTCTT 1680
DB      488 AspLysValIleValAlaPheValSerArgLysAlaValAlaAspHisLeuSerSerAspLeu 507
QY      1681 TGTCAGAAAAAATATCAACTACAGATTCATGATGATGCGGACAGAGAGCGGAG 1740
DB      508 lIeLeuGlyAsnIleSerValGluSerLeuHisGlyAspArgGluGlnArgAspArgGlu 527
QY      1741 CAAGCTCTTGAGACTTTTGGCTTGAAGTGCCAGCTTCTTGTGCTACTTACATAGT 1800
DB      528 LysAlaLeuGluAsnPheLysThrGlyLysValArgIleLeuIleAlaThrAspLeuAla 547
QY      1801 GCCAGAGGGCTGATATTTGAAATGTGCACAACATGTTATCAATTTTACTCTTCTTACC 1860
DB      548 SerArgLysLeuAspValHisAspValThrHisValTyrAsnPheAspProArgAsn 567
QY      1861 ATTGATGAATATGCTCATGCAATTTGGCGCTACTGTCGTGTTGCGAATACTGCGACAGCA 1920
DB      568 lIeGluGluTyrValHisArgIleGlyArgThrGlyArgAlaGlyArgThrGlyValSer 587
QY      1921 ATTTCCTTTTGTGATTTGATCGGATTAACATTAGACACAGCTCTAGTAAAGATTGG 1980
DB      588 lIeThrThrLeu---ThrArgAsnAspTyrArgValAlaSerGluLeuIleAsnIleLeu 606

QY      1981 ACAGATGCTCAACAGATGTTCTCGATGCTGGAAGAAATTGCC 2025
DB      607 GUARGAlaasnGInserileProGluGluLeuValserMetala 621

RESULT 7
US-09-208-742-4
; Sequence 4, Application US/09208742
; Patent No. 6174679
; GENERAL INFORMATION:
; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: CIP150/HTAFI150 is Necessary for Cell
; FILE REFERENCE: 1453.002
; CURRENT APPLICATION NUMBER: US/09/208,742
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: human
US-09-208-742-4

Alignment Scores:
Pred. No.:      1,39e-64      Length:      1261
Score:          704.00      Matches:      227
Percent Similarity: 41.98%      Conservative: 113
Best Local Similarity: 28.02%      Mismatches: 264
Query Match:    17.80%      Indels:      206
DB:             4           Gaps:          26

US-09-714-865-15 (1-2172) x US-09-208-742-4 (1-1261)

QY      139 GGAAGCTTCTGGAAGATCATTTTCATGAAGATTTGCTTGGCGGGAATTTTGA 198

```





```

Db      486 PheArgLysValIleGluLysLeuAlaArgAspIleLeu-----IleAspProIleArgVal 504
Qy      1510 GCTGTGGACAGATGGGTGGACATGTAGAGATGTTGACAGACCGTT-----CTCCAA 1563
Db      505 ValGlnGlyAspIleGlyGluAlaAsnGluAspValThrGlnIleValGluIleLeuHis 524
Qy      1564 GTTGGCCAGTTCCTCAAA-----AGAGAAAGCTCGTTGAAATTCGCGAAAC 1611
Db      525 SerGlyPro--SerLysTrpAsnTrpLeuThrArgArgLeuValGluPhe----- 540
Qy      1612 ATAGGGGATGAAGAAGATAGTGTCTTGTGTAACATAAGAAAAAGCAGATTTTACTGCA 1671
Db      541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnIleGluIleValAla 560
Qy      1672 ACTTTTCTTTGTCAAGAAAAATATCAACTACAGTATCCATGTGTGTCGGGAACAGAGA 1731
Db      561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580
Qy      1732 GAGCGGAGCAGAGCTGTGGAGATTTTGGCTTGGAAAGGCCAGATTCCTTGTGTGCTACT 1791
Db      581 GluArgAsnLysValIleSerAspPheLysLysAspIleProValLeuValAlaThr 600
Qy      1792 TCAGTACCTCCAGAGCGGTGATATTAATAATGTCACACATGTTATCAATTTTGTATCTT 1851
Db      601 AspValAlaIleAlaArgLysLeuAspIleProSerIleLysThrValIleAsnTyArgPheVal 620
Qy      1852 CCTTACCATGATGATATGTTTCATGCAATTTGGCGTACTGCTGCTGGCAATACT 1911
Db      621 AlaArgAspIleAspTrpIleAsnThrHisArgIleGlyArgThrGlyArgAlaGlyLys 640
Qy      1912 GGCAGACCAATTCCTTTTGTGATCTTGATTCGATACGATTAACCATTAAGACAGCCCTAGTA 1971
Db      641 GlyValAlaLysTrpThrLeuLeu--ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659
Qy      1972 AAGTATATGACAGATGCTCAACAGATGTTCTT----- 2004
Db      660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGlyLeuLeuAspLeuAlaMetGln 679
Qy      2005 --GCATGGTTGAGAAATATGCTTTAGT----- 2031
Db      680 AsnAlaTrpPheArgLysSerArgPheLysGlyGlyLysGlyLysLeuAsnIleGly 699
Qy      2032 -----ACATACATTCCTGGCTTCAGTGTAGTACAA-----AGAGCA 2067
Db      700 GlyGlyLysLeuGlyTyArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719
Qy      2068 AAC----- 2070
Db      720 AsnAsnAsnValMetSerAsnTyArgLysAlaLysTrpProSerThrGlyAlaMetGlyAsp 739
Qy      2071 -----GTTTTCGCA 2079
Db      740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTyLysSerHisPheValAlaAla 759
Qy      2080 TCAGTGTATACCAAGAAG--GGCAGAGACCTTGAACACAGCTGGGTTTCTTCTCA 2136
Db      760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTTrpHisSerAla 779
Qy      2137 CGAGCTCCCAATCCAGATGATGATGATCA 2166
Db      780 GlySerLeuAsnSerValProThrAsnSer 789

```

```

; PRIOR APPLICATION NUMBER: US 09/332,295
; PRIOR FILING DATE: 1999-06-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1261
; TYPE: PR
; ORGANISM: Homo sapien
US-09-709-979-2

Alignment Scores:
Pred. No.: 1,39e-64 Length: 1261
Score: 704.00 Matches: 227
Percent Similarity: 41.98% Conservative: 113
Best Local Similarity: 28.02% Mismatches: 264
Query Match: 17,80% Indels: 206
DB: 4 Gaps: 26

US-09-714-865-15 (1-2172) x US-09-709-979-2 (1-1261)
Qy      139 GGCACCTTCGAAAGATCATTTGATGAAGAAAGTGTGCTGCGCGAATTTTGA 198
Db      52 GlyProGlyThrLysArgGlyPheGlyPheGlyGlyPheAlaIle-----SerAlaGly 69
Qy      199 AACAGAGATGCTGTGAGTGTATTAAGCAGATTAAT-----ACATCCACA 243
Db      70 LysLysGluGluProLysLeuProGlnGlnSerHisSerAlaPheGlyAlaThrSerSer 89
Qy      244 ATGGGTGTTTGA-----GTTGAAAG 267
Db      90 SerSerGlyPheGlyLysSerAlaProProGlnLeuProSerPheTyLysIleGlySer 109
Qy      268 AGTTTGAAGAACAGAGCTTTTCAACACAGAGTTTGAAGTGTAGTGTGCTGTTTC 327
Db      110 LysArgAlaAsnPheAspGluGluAsnAlaTyPheGluAspGluGlu----- 125
Qy      328 TGGAGAGCTTATTAAT-----GACTCGAAGATTAATCCACACCG 369
Db      126 ---GluAspSerSerAsnValAspLeuProTyLleProAlaGluAsnSerProThrArg 144
Qy      370 AACAGAGGCTTTCCAGAGAGCGGCTATCCAGATGAAATTAATTCAGAGCTTACGG 429
Db      145 GlnGlnPheHisSerLys----- 150
Qy      430 CCATACAGAAAGAGGTGAAGAGTACTTCCAGAGTTGCCGTGAGAGATTGGCTTAGCA 489
Db      150 ----- 150
Qy      490 AGTCAATATATGACTTAGACCCAGACGAATGATGACGACCTGTGGCTTTTGGT 549
Db      151 ProValAspSerSerAspAspAsp----- 159
Qy      550 TTAGAAGACCAATTAAGTGCACAGATATGATGATCTTCTCAAGCAGA----- 603
Db      160 -----ProLeuGluAlaPheMetAlaGluValGluAspGlnAlaAlaArgAspMet 176
Qy      604 -----AGTGCAGTGAAGTGAACGAGCGTGTACCAAGGTTTAATGAAGAGTA 654
Db      177 LysArgLeuGluGluLysAspLysGluArgLysAsnValLysGlyLysLeuArgAspPheIle 196
Qy      655 ATAACAGGCTCTTGAAAGAATTTCTGC----- 681
Db      197 GluGluGluAspAspGlnGluAlaTyPheArgTyPheMetAlaGluAsnProThrAlaGly 216
Qy      682 -----AAGTCAGACAGCAGAGAGAGAGAAAGTAGTATCTCAAGCA----- 723
Db      217 ValValGlnGluGluGluGluAsnLeuGluTyArgAspSerAspGlyAsnProIleAla 236
Qy      724 CCAAAAGTACCTACATACCCCTCTCCACCTCGAGAT----- 762
Db      237 ProThrLysLysIleIleAspProLeuProIleAspHisSerGluIleAspTyPro 256
Qy      763 -----GAGGACTCCACTTTGCACATTAACAGACAGCATAAAC----- 801

```

```

RESULT 9
; Sequence 2, Application US/09709979
; Patent No. 6423822
; GENERAL INFORMATION:
; APPLICANT: Kautmann, Joerg
; TITLE OF INVENTION: CIP130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: 200130.456 / 1513.003
; CURRENT APPLICATION NUMBER: US/09/709,979
; CURRENT FILING DATE: 2000-11-09

```

```

Db      257 ProPheGluYsaSnPheTYrAsnGluHISgluGluIleThrAsnLeuThrProGln 276
QY      802 ---TTCCAGAAATACGACACTATCTTGAGAGTGTGGACATGATGCACACGACA 858
Db      277 LeuIleAspLeuAlaGHisLysLeuAsnLeuAlaValSerGlyAlaAlaProProArgPro 296
QY      859 ATTTCGACTTTTGAAGAAGCTAATCTCTGTGCAGACTGAATMAACAATTGCTAAAGCT 918
Db      297 GlySerSerPheAlaHisPheGlyPheAspGluGlnLeuMetHISglnIleArgLysSer 316
QY      919 GGTATATCTAAGCTTACTCTCTGTGCAAAAATACAGTATCTTATCATATCTTGCGAGCA 978
Db      317 GluTYrThrGlnProThrProIleGlnCysGlnIleValAlaLeuSerGlyArg 336
QY      979 GATTATGCGCTTGTGTCACAAACAGGCTGGGAAGACTGGCGCTTTCTCTTACCAATT 1038
Db      337 AspmetIleGlyIleAlaLysThrGlySerGlyThrAlaAlaPheIleTrpProMet 356
QY      1039 TTGGCTCATATGATGATGATGAATACTGCCAGTCGTTTAAAGAGTTCACAGAA-- 1095
Db      357 LeuIleHisIleMetAspGln-----LysGluLeuGlnProGly 369
QY      1096 -----CCAGAGTATATTATTTGTGACCAACTGCAGAAATTGTCACACAGATTATTG 1149
Db      370 AspGlyProIleAlaValAlaLeuValCysProThrArgLeuLeuCysGlnGlnIleHisPro 389
QY      1150 GAAGCCAGAAATTTCTTTTGGGACTTGTGTAAGCTGTTGTTATATATATGGGGGAACC 1209
Db      390 GluYsaLysAspPheGlyLysAlaTYrAsnLeuArgSerAlaAlaValTYrGlyGlyGly 409
QY      1210 CAGCTGGACATTCATTCGACAAATAGTACAGGCTGTATATATTATGTCATCTCT 1269
Db      410 SermetTrpGluGlnAlaLysAlaLeuGlnGluGlyAlaGluIleValValCysThrPro 429
QY      1270 GGAAGACTGATGATATCATATGCGAAAGAAAGTTGGTCTTCAACACAGATCAAAATCTTA 1339
Db      430 GlyArgLeuIleAspHisValLysLysAlaThrAsnLeuGlnArgValSerTYrLeu 449
QY      1330 GTTTTGATGAAGCTGATCGCATGCTGATGCTGGTATGGTTTGTCCAGAAATGAAGATTA 1389
Db      450 ValPheAspGluAlaAspArgMetPheAspmetGlyPheGluTYrGlnValArgSerIle 469
QY      1390 ATTTCCTGCCAGGAAATGCCATCAAGGACAGCGCCAAACCTTATGTTGATGTCAGTCAACT 1449
Db      470 Alaser-----HisValArgProAspArgGlnThrLeuLeuPheSerAlaThr 485
QY      1450 TTTCGACAGAAATTCAAAGGTTGGCTGCAGAGTTTAAAGTCAAAATTCATCTGTTGTT 1509
Db      486 PheArgLysLysIleGluLysLeuAlaArgSpIleLeu--IleAspProIleArgVal 504
QY      1510 GCTGTGGACAAAGGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
Db      505 ValGlnIleAspIleGlyGluAlaAsnGluAspValThrGlnIleValGlnIleLeuHis 524
QY      1564 GTTGGCCAGTTCTCAAAA-----AGAGAAAAGCTCGTGAATTCCTCGGAAC 1611
Db      525 SerGlyPro---SerLysTrpAsnTrpLeuThrArgArgLeuValGluPhe----- 540
QY      1612 ATAGGGGATGAAGAACTAGTGTCTTTGTTGAAAATAAGAAAAAGACAGATTTTACTGCA 1671
Db      541 ThrSerSerGlySerValLeuLeuPheValThrLysLysAlaAsnAlaGluLeuAla 560
QY      1672 ACTTTTCTTTGTCAGAGAAAAATATCAACTCAAGTATCCATGCTGATCGGGAACAGAGA 1731
Db      561 AsnAsnLeuLysGlnGluGlyHisAsnLeuGlyLeuLeuHisGlyAspMetAspGlnSer 580
QY      1732 GAGCGGAGACAGCTCTTGAGATTTTGTGCTTGGAAAGTCCCGAGTTCTGTGCTACT 1791
Db      581 GluArgAsnLysValIleSerAspPheLysLysAspIleProValLeuAlaAlaThr 600
QY      1792 TCAGTAGCTGCAGAGGCTGATATTGAATAATGTGCAACATGTTATCAATTTTGATCTT 1851

```

```

Db      601 AspValAlaAlaArgGlyLeuAspIleProSerIleLysThrValIleAsnTYrAspVal 620
QY      1852 CCTTCTACCATTTGATGATATGTTTCATGCAATTTGGCGTACTGCTGTTGGGAATTA 1911
Db      621 AlaArgAspIleAspTrpHisThrHisArgIleGlyArgThrGlyArgAlaGlyGluLys 640
QY      1912 GGCAGACCAATTTCTTTTGTGATCGGAATCGGAATCAACATTTAGACAGCTCTAGTA 1971
Db      641 GlyValAlaTYrThrLeuLeu---ThrProLysAspSerAsnPheAlaGlyAspLeuVal 659
QY      1972 AAAGTATTCAGACAGTGTCTCAACAGAGATGTTCTCT----- 2004
Db      660 ArgAsnLeuGluGlyAlaAsnGlnHisValSerLysGluLeuAspLeuAlaMetGln 679
QY      2005 ---GCATGTTTGAAGAATTTGCCCTTAGT----- 2031
Db      680 AsnAlaTrpPheArgLysSerArgPheLysGlyGlyArgGlyLysLysLeuAsnIleGly 699
QY      2032 -----ACATACATTCCTGCTTCAGTGTAGTACA-----AGAGCA 2067
Db      700 GlyGlyLysLeuGlyTYrArgGluArgProGlyLeuGlySerGluAsnMetAspArgGly 719
QY      2068 AAC----- 2070
Db      720 AsnAsnAsnValMetSerAsnTYrGluAlaTYrLysProSerThrGlyAlaMetGlyAsp 739
QY      2071 -----GTGTTGCA 2079
Db      740 ArgLeuThrAlaMetLysAlaAlaPheGlnSerGlnTYrLysSerHisPheValAlaAla 759
QY      2080 TCAGTGTATACCAAGAG---GGCAGAGACACTTTGAACACAGCTGGGTTTCTTTCTTCA 2136
Db      760 SerLeuSerAsnGlnLysAlaGlySerSerAlaAlaGlyAlaSerGlyTrpThrSerAla 779
QY      2137 CGAGCTCCCAATCCAGTATGATGATGATCA 2166
Db      780 GlySerLeuAsnSerValProThrAsnSer 789

RESULT 10
US-09-134-001C-4262
; Sequence 4262, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4262
; LENGTH: 528
; TYPE: PRN
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4262

Alignment Scores:
Pred. No.: 2,036-50
Score: 566.50
Percent Similarity: 53.23%
Best Local Similarity: 33.83%
Query Match: 14.33%
DB: 4
Caps: 8
US-09-714-865-15 (1-2172) x US-09-134-001C-4262 (1-528)

QY      868 TTGAAGAGCTATCTCTGTGCAGACCTGAATTAACAATGTGTAAGCTGTTATACT 927
Db      23 PheLysGluLeuGlyIleSerAspLysThrValGlnThrLeuGluAlaMetGlyPheLys 42

```

```

QY 928 AAGCTTACTCTGTGCAAAAATACAGTATTCCTCATCTGACGAGACGATTGATG 987
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43 GluProthrProIleGlnLysAspSerIleProIleValLeuGlnLysAspIleLeu 62
QY 988 GCTTGCTCTCAACAGAGCTGTGGAGAGCTGCGCTTTCTCCACCAATTTGGCTCAT 1047
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 GlyIleAlaGlnThrGlyThrGlyLysThrGlyLysLeuIleProLeuIleGlnLys 82
QY 1048 ATGATGATGATGATGATTAATTAATCCAGTCTTTAAAGAGTTGCAGAACCAAGCTGATTT 1107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 ValValGlyGln-----GlnGlyValGlnSerLeu 92
QY 1108 ATTGTACCAACCAACTGCAAGATTGTCACACAGATTTATTGGAAGCCAGAAATTTTCT 1167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 IleLeuAlaProthrAlaGlnLeuAlaMetGlnValAlaGlnLysLeuGlnLysPheSer 112
QY 1168 TTGGGACTTGTGTAAAGCTGTTGTTATATATATGAGGGAACCCAGCTGGACATTCATTT 1227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 LysGlyGlnLysValGlnValAlaThrValPheGlyGlyMetProIleGlnLysGlnIle 132
QY 1228 CGACAAATAGTACAGGCTGTAATATATATATGCTACTCTCTGAAAGACTGATGATATC 1287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 LysAlaLeuLysArgGlyProGlnIleValValGlyThrProGlyArgValIleAspHis 152
QY 1288 ATAGGCAAAAGAAAGATTGCTCTCAACAGATCAAAATCTAGTTTGGAGAGCTGAT 1347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 LeuAsnArgThrLeuLysThrGlnLysIleHisThrLeuIleLeuAspGlnAlaAsp 172
QY 1348 CGCATGTTGATGATGGTTTGTGTCAGAAATGAAGAAATTTCTTGCCCAAGATG 1407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 GlnMetMetAsnMetGlyPheIleAspAspMetArgPheIleMetAsp-----LysIle 190
QY 1408 CCATCAAGAAAGACAGCCCAACCTTATGTTGCTAGTCACTTTTCCAGAGAAATTCAA 1467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 ProAla---GlnGlnArgGlnThrMetLeuPheSerAlaThrMetProLysAlaIleGln 209
QY 1468 AGCTTGCTGCAAGATTTTAAAGTCAATATCTGTTGTGCTGTT---GGCAAGTG 1524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 GlnLeuValGlnGlnPheMetLysAlaProLysIleIleLysThrMetAsnAsnGlnLec 229
QY 1525 GGTGAGCATGATAGATGTTCCAGACAGCCGTTCTCCAAAGTTGGCCAGTTCTCAAAAAGA 1584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 SerAspProGlnIleAspGlnLysThrIle-----ValLysGlnLysGlnLysPhe 247
QY 1585 GAAAGCTCGTTGAATTCGCGAAACATAGGGGATGAAGAAATATGCTCTTTGTTGAA 1644
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 AspThrPheThrAsnPheLeuAspValHisGlnProGlnLeuAlaIleValPheGlyArg 267
QY 1645 ACTAAGAAAAGAGCATTTTACTGCACTTTCTTTGTCAGAAAATATCAACTCA 1704
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 ThrLysAspArgValAspGlnLeuThrSerAlaLeuLysSerLysGlyThrLysAlaGln 287
QY 1705 AGTATCCATGCTGATCGGGAACAGAGAGCGGAGCAAGCTCTTGAGATTTTCGCTTT 1764
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GlyLeuHisGlyAspIleThrGlnAlaLysArgLeuGlnValLeuLysLysPheLysAsn 307
QY 1765 GGAAGTCCCAAGTTCTTGTGCTACTTCACTAGTCTGCCAGAGGCTGGATATGAAAT 1824
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 AspGlnIleAspIleLeuValAlaThrAspValAlaAlaArgGlyLeuAspIleSerGly 327
QY 1825 GTGCAACATGTTATCAATTTGATCTTCTTACCATTTGATGATATGTCATGATTT 1884
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 ValSerHisValThrAsnPheAspIleProGlnAspThrGlnLysSerThrHisArgIle 347
QY 1885 GGGCGTACTGCTGCTGGGAAATCTGCGAGAGCAATTTCTTTTGTAT----- 1935
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 GlyArgThrGlyArgAlaGlyLysGlnLysIleAlaValThrPheValAsnProIleGln 367
QY 1936 -----CTGATCGGATTAACAT-----TTACCAAG 1962
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 MetAspTyrIleArgGlnIleGlnLysAspValAsnAsnArgArgMetLysAlaLeuArgPro 387

```

```

QY 1963 CCTTAGTAAAGATATGACAGATGCTCAACAGAT-----GTTCTGCA 2007
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 ProHisArgLysLeuValLeuLysValArgGlnAspAspIleLysAspArgValGlnAsn 407
QY 2008 TGGTTG 2013
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 TrpMet 409

RESULT 11
US-09-318-443-8
; Sequence 8, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Bivandjou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-318-443-8

Alignment Scores:
Pred. No.: 4,27e-44 Length: 411
Score: 505.50 Matches: 130
Percent Similarity: 51.38% Conservative: 93
Best Local Similarity: 29.95% Mismatches: 150
Query Match: 12.78% Indels: 61
DB: 4 Gaps: 10

US-09-714-865-15 (1-2172) x US-09-318-443-8 (1-411)
QY 757 GAGATGAGATGATCTTCATCTTGCACTTATTCAGACGACATTAACCTGCAATAACGAC 816
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 LysGlnLysAsp-----MetThrLysValGlnPheGlnThrSerGln 32
QY 817 ACTATCTTGTGGAAGTCTGACATGATGACCAACCAATTCGATTTGAGAA 876
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 GluVal-----AspValThrPro-----ThrPheAspThr 42
QY 877 GCTATCTCTGTCAGACATGATTAACACATTCGTAAGCTGTATTAAGCTTACT 936
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 43 MetGlyLeuArgGlnAspLeuLeuArgLysIleThrLysArgLysPheGlnLysProSer 62
QY 937 CCGTGGAAAATACATATTCATCATCTTGCAGAGAGAGATTGATGGCTTGCT 996
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 AlaIleGlnGlnArgAlaIleLysGlnIleLysGlyLysArgAspValIleAlaGlnSer 82
QY 997 CAACAGGCTCTGGGAAGACTGCGGCTTTCTCTTACCAATTTGGCTCATATGATGCAT 1056
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 GlnSerGlyThrGlyLysThrAlaThrPheSerIleSerValLeuGlnLysLeu----- 100
QY 1057 GATGGAATTAACCTGCCAGCTGTTTAAAGACTGCGAGAACACAGCTGATATTTGTCGA 1116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 101 -----AspIleGlnValArgGlnThrGlnAlaLeuIleLeuAla 113
QY 1117 CCAACTGAGAAATTTGTCACACGATTTATTGGAACCGAAATTTCTTTGGGACT 1176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 ProThrArgGlnLeuAlaValGlnIleGlnLysGlyLeuLeuAlaLeuGlyAspTyrMet 133
QY 1177 TGTGTAGAGCTGTTGTTATATATGAGGGAACCCAGCTGGACATTCATTTGCAATA 1236
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 AsnValGlnCysHisAlaCysIleGlyLysThrAsnValGlyGlnAspIleArgLysLeu 153
QY 1237 GTACAGAGCTGTATATATATATGCTACTCTCGAAGACTGATGATATCATGCGCAA 1296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 AspTyrGlyGlnHisValValAlaGlyThrProGlyLysValPheAspMetIleArgArg 173

```



```

QY 1297 GAAAAAGATTGCTCTCAACAGATCAATTAAGTATTTGATGAGTGTGATGCGATGTTG 1356
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 174 ArgSerLeuArgThrArgAlaIleLeuMetLeuValLeuAspGluAlaAspGluMetLeu 193
QY 1357 GATATGGCTTTTGGTCCAGAAATGAGAAAGTTAATTTCTTCCAGGAATGCCATCAAG 1416
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 194 AsnLysGlyPhe-----Lys 198
QY 1417 GAACAG-----CGCAACCCCTTAATGTCAGT 1443
    |||||
Db 199 GluGlnIleTyrAspValTyrArgTyrLeuProProAlaThrGlnValValLeuIleSer 218
QY 1444 GCAACCTTTCCAGAGAAATTCAGAGTGTGCTGAGATTTTAAAGTCAATATATCTG 1503
    |||||
Db 219 AlaThrLeuProHISGluIleLeuGluMetThrAsnLysPheMet---ThrAspProIle 237
QY 1504 TTTCCTGCTGTGGACAAGTGGGTGAGCATGTAGAGATGTTCAAGACCGTCTCCAA 1563
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 238 ArgIleLeuValLysArgAspGluLeuThrLeuGluGlyIleLysGlnPhePheValAla 257
QY 1564 GTTGCCAGTTCCTCA-----AAAAGAAAAGCTGCTGAATTCGCCAAACATAGGGGAT 1620
    |||||
Db 258 ValGluArgGluGluTyrPlysPheAspThrLeuLysAspLeuTyrAspThrLeuThrIle 277
QY 1621 GAAAGAACTATGCTTTGTTGAACCTAAGAAAAGAGAGATTTTACTGCAACTTTCTT 1680
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 278 ThrGlnAlaValIlePheCysAsnThrLysArgLysValAspThrPleuThrGluLysMet 297
QY 1681 TGTCAAGAAAATATATCACTACAGATATCCATGCTGATCGGGAACAGAGAGCGGAG 1740
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 298 ArgGluAlaAsnPheThrValSerSerMetHisGlyAspMetProGlnLysGluArgGlu 317
QY 1741 CAAGCTCTGGAGATTTTCGCTTTGSAAGTCCAGTCTGTTGCTACTTCACTGAGT 1800
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 318 SerIleLeuLysGluPheArgSerGlyAlaSerArgValLeuIleSerThrAspValTyr 337
QY 1801 GCCAGAGGCTGATATATGAAATATGTCACATGTTATCAATTTGATCTTCTTCAAC 1860
    |||||
Db 338 AlaArgLysLeuAspValProGlnValSerLeuIleIleAsnTyrAspLeuProAsnAsn 357
QY 1861 ATTGATGAATATGTTCAATCGAATTTGGCGGTACTGCTGTTGGGAATACTGGCAGACA 1920
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 358 ArgGluLeuTyrIleHisArgIleGlyArgSerGlyArgTyrGlyArgLysGlyValAla 377
QY 1921 ATTTCCTTTTGTATCTTGAATCGGATTAACATTTAGACAGCCTTAGTAAAGTATG 1980
    |||||
Db 378 IleAsnHe-----ValLysAsnAspAsp-----IleArgIleLeu 389
QY 1981 ACAGATGCTCAACAGATGTTCTCGCATGTTGGAAGAATT 2022
    |||||
Db 390 ArgAspIleGluGlnTyrTyrSerThrGlnIleAspGluMet 403

RESULT 12
US-09-318-443-2
; Sequence 2, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyvanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-09-318-443-2

```

```

Alignment Scores:
Pred. No.: 5,48e-44 Length: 415
Score: 504.50 Matches: 131
Percent Similarity: 50.81% Conservative: 89
Best Local Similarity: 30.25% Mismatches: 152
Query Match: 12.76% Indels: 61
DB: 4 Gaps: 10

US-09-714-865-15 (1-2172) x US-09-318-443-2 (1-415)

QY 760 GATGAGACTTCATCTTTCGACATTAACAGACAGGCAATAACTTCGACAAATAGCACACT 819
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 23 GluGlnLysPheThrThrValGluPheGlnThrSer----- 34
QY 820 ATTCTTGGAAGTGTGGACATGATGACCAACACCAATTCGACTTTGAAGAGCT 879
    |||||
Db 35 -----GluGlnVal-----AspValThrPro-----ThrPheAspThrMet 46
QY 880 AATCTGTGACACACTGAATTAACACATTTGCTAAAGCTGTTATTAAGTCTTACTCT 939
    |||||
Db 47 GlyLeuArgGluAspLeuLeuArgGlyIleTyrAlaTyrGlyPheGluLysProSerAla 66
QY 940 GTGCAAAAATACACTATTCCTATCATCTTGACAGACAGAGATTTGATGCTTGCTCAA 999
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 67 IleGlnGlnLysAlaIleLysGlnIleLysGlyArgAspValIleAlaGlnSerGln 86
QY 1000 ACAGGCTGTGGAAAGTGGGCTTTCTCTCCATTTTGGCTCATATGATGATGATGAT 1059
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 87 SerGlyThrGlyLysThrAlaThrPheCysAlaSerValLeuGlnCysLeu----- 103
QY 1060 GGAATTAAGTCCAGTCTTTTAAAGAGTGTGACAGACCAAGTATATTTGATGACACA 1119
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 104 -----AspIleGlnIleArgGluThrGlnAlaLeuIleLeuAlaPro 117
QY 1120 ACTCGAATTTGTCACCAAGATTTATTTGGAAGCCAGAAATTTCTTTGGGACTTGT 1179
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 118 ThrLysGluLeuAlaArgGlnIleGlnLysValLeuLeuAlaLeuGlyAspTyrMetAsn 137
QY 1180 GTAAGACTGTGTTATATATGAGGGGAACCCAGCTGGACCTTCAATTCGACAAATGTA 1239
    |||||
Db 138 ValGlnCysHisAlaCysIleGlyGlyThrAsnValGlyLysAspIleArgLysLeuAsp 157
QY 1240 CAAGGCTGTATATATATATGCTGCTACTCTGGAAGCTGATGATCATATGAGCAAGAA 1299
    |||||
Db 158 TyrGlyGlnHisValValAlaIleGlyThrProGlyArgValAlaPheAspMetIleArgArgArg 177
QY 1300 AAGATGCTCTCAACAGATCAATTAATGTTGATGATGATGATGATGATGATGAT 1359
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 178 SerLeuArgThrArgAlaIleLysMetLeuValLeuAspGluAlaAspGluMetLeuAsn 197
QY 1360 ATGGGTTTGGTCCAGAAATGAGAAAGTAAATTTCTTCCAGGAATGCCATCAAAGGA 1419
    |||||
Db 198 LysGlyPhe-----LysGlu 202
QY 1420 CAG-----CGCAACCCCTTAATGTCAGTCA 1446
    |||||
Db 203 GluIleTyrAspValTyrArgTyrLeuProProAlaThrGlnValCysLeuIleSerAla 222
QY 1447 ACTTTCCAGAGAAATTCAGAGTGTGCTGACAGATTTTAAAGTCAATATATCTGTT 1506
    |||||
Db 223 ThrLeuProHISGluIleLeuGluMetThrAsnLysPheMet---ThrAspProIleArg 241
QY 1507 GTTGCTTTGGACAAGTGGGTGAGCATGTAGAGATTTTCAAGCAGACCGTCTTCAAGTT 1566
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 242 IleLeuValLysArgAspGluLeuThrLeuGluGlyIleLysGlnPhePheValAlaVal 261
QY 1567 GGCAGTTCCTCA-----AAAAGAAAAGCTGCTGAATTCGCCAAACATAGGGGATGAA 1623
    |||||
Db 262 GluArgGluGluTyrPlysPheAspThrLeuLysAspLeuTyrAspThrLeuThrIleThr 281
QY 1624 AGAATATGCTTTGTTGAACCTAAGAAAAGCAGATTTTACTGCAACTTTCTTTGT 1683
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 282 GluAlaValIlePheCysAsnThrLysArgLysValAspThrPleuThrGluLysMetArg 301

```

```

QY 1684 CAAGAAAAATATCAACTACAGATATCATGTGATCGGAAACAGAGACGGGAGCAA 1743
   ::::::::::::::::::::|
Db 302 GJUALAAsphenrthValSerSerMetHsclYAspMetProGlnHvsgluHArgGluSer 321
QY 1744 GCTCTTGAGATTTTGGCTTTGGAAAGTCCAGTCTTGTGTCTTACTTCAAGTCC 1803
   ::::::::::::::::::::|
Db 322 IlleMetLysGluPheArgSerGlyAlaSerArgValIleuIleSerThrAspValIlePheAla 341
QY 1804 AGAGGCGCTGATTTGAAATATGTCACATATTTGATTTGATTTCTTCTTACCAT 1863
   |||||
Db 342 ArgGlyLeuAspValProGlnValSerIleIleAsnTyraAspLeuProAsnAsnArg 361
QY 1864 GATGAATATGTCATTCGATTTGGGCGTACTGCTGTTGGGAAATPACTGGCAGAGCAATT 1923
   ::::::::::::::::::::|
Db 362 GluLeuTyrlleHsArgGlyIleGlyArgSerGlyArgGlyValAlaIle 381
QY 1924 TCCCTTTTGTGATCTGGAATCGGATTAACCATTTAGCACAGCCTTAGTAAAGTATTGACA 1983
   ::::::::::::::::::::|
Db 382 AsnPhe-----ValLysAsnAspAsp-----IleArgIleLeuArg 393
QY 1984 GATGCTCAACAGATGTTCTCGCATGTTGGAAGAAATT 2022
   |||||
Db 394 AspIleGluGlnTyrlYrSerThrGlnIleAspGluMet 406

RESULT 13
US-09-318-443-6
; Sequence 6, Application US/09318443
; Patent No. 6197947
; GENERAL INFORMATION:
; APPLICANT: Hemmacti-BriVanlou, Ali
; APPLICANT: Weinstein, Daniel C.
; TITLE OF INVENTION: TRANSLATION INITIATION FACTOR 4AIII, AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-211 N
; CURRENT APPLICATION NUMBER: US/09/318,443
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-318-443-6

Alignment Scores:
Pred. No.: 6,93e-44 Length: 411
Score: 503.50 Matches: 128
Percent Similarity: 53.43% Conservative: 98
Best Local Similarity: 30.26% Mismatches: 158
Query Match: 12.73% Indels: 39
Gaps: 10

US-09-714-865-15 (1-2172) x US-09-318-443-6 (1-411)
QY 757 GAGATGAGGAGCTCATCTTTGACATATTCAGACAGCATTAACCTTGACAAATATAGAC 816
   ::::::::::::::::::::|
Db 19 LysGluGluAsp-----MetThrLysValGluPheGlnThrSerGlu 32
QY 817 ACTATCTTGTGGAAGTCTGACATGATGACACACACGCAATTCGATCTTTGAGAA 876
   ::::::::::::::::::::|
Db 33 GluVal-----AspValThrPro-----ThrPheAspThr 42
QY 877 GCTATCTCTGTCAGACACTGATATACAACTGTGTAAGCTGTTATATCAAGCTTACT 936
   |||||
Db 43 MetGlyLeuArgGluAspLeuLeuArgGlyIleTyraIleTyrlYrGlyPheGlnLysProSer 62
QY 937 CCTGGCAAAATACAGTATTCATCATCTGATGAGAGAGATTGATGAGCTTGTGCT 996
   ::::::::::::::::::::|
Db 63 AlaIleGlnGlnAlaIleLysGlnIleIleLysGlyArgAspValIleAlaGlnSer 82
QY 997 CAACAGGCTCTGGAGAACTGCGCTTTTCTCTACCAATTTGGCTCATATATGATGCAT 1056
   |||||

```

```

Db 83 GlnSerGlyThrGlySerThrAlaThrPheSerIleSerValLeuGlnCysLeu----- 100
QY 1057 GATGAATAACTGCGACGCTTTTAAAGATGTCAGACACAGAGTATATATATGATGCA 1116
   ::::::::::::::::::::|
Db 101 -----AspIleGlnValArgGluThrGlnAlaIleLeuAla 113
QY 1117 CCAATGAGAAATTTGTCACACAGATTTATTTGGAACCCAGAAATTTTCTTTGGAGCT 1176
   |||||
Db 114 ProThrArgGluLeuAlaValGlnIleGlnLysGlyLeuLeuAlaLeuGlyAspTyrlMet 133
QY 1177 TGTGTAAAGACTGTTGTTATATATATGAGGGAACCCAGCTGGACATTTATTCGACAAATA 1226
   |||||
Db 134 AsnValGlnCysHsAlaCysIleGlyGlyThrAsnValGlyIleAspIleArgLysLeu 153
QY 1237 GTACAAAGCTGTATATATTATGCTACTCTCGAAGACTGATGATCATCATGCGCAA 1296
   |||||
Db 154 AspTyrlGlyGlnHsValValAlaGlyThrProGlyArgValPheAspMetIleArgArg 173
QY 1297 GAAAAGATGCTGTCAACAGATCAATCAATCTAGTTTGGATGAGCTGATCGCATGTTG 1356
   ::::::::::::::::::::|
Db 174 ArgSerLeuArgThrArgAlaIleLysMetLeuValLeuAspGluAlaAspGluMetLeu 193
QY 1357 GATATGGCTTTTGGTCCAGAAATGAGACAGTAAATTTCTTGCCAGAAATCCCATCAAG 1416
   ::::::::::::::::::::|
Db 194 AsnLysGlyPheLysGluGlnIleTyraAspValTyraArgTyrl-----LeuProSer--- 210
QY 1417 GAACAGCGCAACCCCTATGTTCACTGCAACCTTTGTCAGAGAAATTCAGAGTTGGCT 1476
   |||||
Db 211 ---AlaThrGlnValValIleIleSerAlaThrLeuProHsGluIleLeuGluMetThr 229
QY 1477 GCAGAGCTTTTAAAGTCAATATCTGTTGCTGCTGGTGGAGAGGCTGAGCATGT 1536
   ::::::::::::::::::::|
Db 230 AsnLysPheMet---ThrAspProIleArgIleLeuValLysArgAspGluLeuThrLeu 248
QY 1537 AGAGATTTGACAGACAGCTTCTTCAAGTTGGCCAGTTCTCA---AAAAGAGAAACCTC 1553
   ::::::::::::::::::::|
Db 249 GluGlyIleLysGlnPhePheValAlaValAlaGluGluGlnTyrlPheAspThrLeu 268
QY 1594 GTTGAATTTGCGGAACATAGAGGAGTGAAGAAGTATGGCTTTGTAACCTAGAGAA 1653
   ::::::::::::::::::::|
Db 269 CysAspLeuTyrlAspThrLeuThrIleThrGlnAlaValIlePheCysAsnThrLysArg 288
QY 1654 AAAGCAGATTTTACTGCACTTTTCTTGTCAAGAAAAATATCACTACATGATTCAT 1713
   |||||
Db 289 LysValAspThrPheLeuThrGluLysMetArgGluAlaAsnPheThrValSerSerMetHs 308
QY 1714 GGTGATCGGGAACAGAGAGCGGAGCAAGCTCTTGAAGATTTTCCCTTTGGAAAGTGC 1773
   |||||
Db 309 GlyAspMetProGlnLysGluArgGluSerIleMetLysGluPheArgSerGlyAlaSer 328
QY 1774 CCAAGTTTGTGCTACTTCAAGTACGTCAGAGGCGCTGATATTTGAAATGTGCACAT 1833
   |||||
Db 329 ArgValLeuIleSerThrAspValThrAlaArgGlyLeuAspValProGlnValSerLeu 348
QY 1834 GTTATCAATTTGATCTTCTTCACTTACCATTTGATGATGATGATGATGATGATGATGAT 1893
   ::::::::::::::::::::|
Db 349 IleIleAsnTyrlAspLeuProAsnAsnArgGluLeuTyrlleHsArgIleGlyArgSer 368
QY 1894 GGTGTTGGGAATPACTGCGACAGCAATTTCTTTTGAATCTTGATCGATACCAT 1953
   |||||
Db 369 GlyGlnTyrlYrGlyArgLysValAlaIleAsnPhe-----ValLysAsnAspAsp--- 385
QY 1954 TTAGCACAGCCTCAGTAAAGTATTGACACAGATGCTCAACAGATGTTCTCGCATGCTTG 2013
   ::::::::::::::::::::|
Db 386 -----IleArgIleLeuArgAspIleGluGlnTyrlYrSerThrGlnIle 400
QY 2014 GAAGAAATT 2022
   ::::::::::::::::::::|
Db 401 AspGluMet 403

RESULT 14
US-08-607-509-4
; Sequence 4, Application US/08607509

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 07:51:10 ; Search time 130 Seconds  
(without alignments)  
5123.861 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 2172  
Sequence: 1 atgagggatgaagatcgga.....tagatgatgatcatcggaat 2172

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued\_Patents\_MA.\*  
2: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	383.2	17.6	2319	US-09-058-489-90	Sequence 90, Appl
2	383.2	17.6	4416	US-09-058-489-17	Sequence 17, Appl
3	369.4	17.0	3408	US-09-058-489-14	Sequence 14, Appl
4	369.4	17.0	5322	US-09-058-489-13	Sequence 13, Appl
5	158.2	7.3	2365	US-09-183-706-42	Sequence 42, Appl
6	158.2	7.3	2365	US-09-567-995-42	Sequence 42, Appl
7	123.6	5.7	1254	PCT-US96-05320A-894	Sequence 84, Appl
8	116.8	5.4	3230	US-08-961-527-203	Sequence 203, Appl
9	115.2	5.3	3825	US-09-208-742-3	Sequence 1425, Appl
10	103.2	4.8	1587	US-09-134-001C-1425	Sequence 1, Appl
11	92.2	4.2	1245	US-09-134-443-1	Sequence 1, Appl
12	91.2	4.2	1368	US-09-134-001C-816	Sequence 816, Appl
13	90.8	4.2	490	US-09-222-575-126	Sequence 126, Appl
14	84.6	3.9	1347	US-08-959-749-1	Sequence 1, Appl
15	84.6	3.9	1347	US-09-351-497-1	Sequence 1, Appl
16	76.2	3.5	1374	US-08-929-738-1	Sequence 1, Appl
17	76.2	3.5	1614	US-08-929-738-2	Sequence 2, Appl
18	73.6	3.4	7218	US-08-232-463-14	Sequence 14, Appl
19	73.6	3.4	1800	US-09-039-773A-1	Sequence 1, Appl
20	72.8	3.4	1682	US-09-318-443-7	Sequence 1, Appl
21	71.2	3.3	1536	US-09-318-443-5	Sequence 1, Appl
22	70.6	3.3	8494	US-08-961-527-163	Sequence 163, Appl
23	65.8	3.0	4774	US-09-221-017B-1	Sequence 1, Appl
24	58.8	2.7	987	US-09-221-017B-1	Sequence 1, Appl
25	58.2	2.7	2934	US-09-149-934-2	Sequence 2, Appl
26	55.2	2.5	43676	US-09-356-952-12	Sequence 12, Appl
27	55	2.5	2061	US-09-008-271A-16	Sequence 16, Appl

28	55	2.5	2061	4	US-09-705-448-2	Sequence 2, Appl
29	53.2	2.4	4403765	4	US-09-103-840A-2	Sequence 2, Appl
30	53.2	2.4	4411529	4	US-09-103-840A-1	Sequence 1, Appl
31	52.8	2.4	1867	2	US-08-607-509-3	Sequence 3, Appl
32	52.8	2.4	1867	3	US-08-634-642-3	Sequence 3, Appl
33	52.8	2.4	1867	3	US-08-989-370-3	Sequence 3, Appl
34	50	2.3	299	4	US-09-118-554-42	Sequence 42, Appl
35	50	2.3	299	4	US-09-118-554-42	Sequence 42, Appl
36	50	2.3	299	4	US-09-602-877A-42	Sequence 42, Appl
37	50	2.3	1380	4	US-09-134-001C-341	Sequence 341, Appl
38	49.6	2.3	1618	2	US-08-533-669A-9	Sequence 9, Appl
39	49.6	2.3	1618	2	US-08-607-509-1	Sequence 1, Appl
40	49.6	2.3	1618	2	US-08-454-036-1	Sequence 1, Appl
41	49.6	2.3	1618	2	US-08-634-642-1	Sequence 1, Appl
42	49.6	2.3	1618	3	US-08-989-370-1	Sequence 1, Appl
43	49.6	2.3	1618	4	US-09-183-861-9	Sequence 9, Appl
44	49.6	2.3	1618	4	US-09-022-765-9	Sequence 9, Appl
45	49.6	2.3	1618	5	PCT-US95-05064-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-058-489-90  
Sequence 90, Application US/09058489  
Patent No. 6103886  
GENERAL INFORMATION:  
APPLICANT: Whitehead Institute for Biomedical Research  
APPLICANT: Lahn, Bruce  
TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of  
FILE REFERENCE: WH197-08PA  
CURRENT APPLICATION NUMBER: US/09/058,489  
CURRENT FILING DATE: 1998-04-10  
EARLIER APPLICATION NUMBER: 60/041,877  
EARLIER FILING DATE: 1997-04-11  
NUMBER OF SEQ ID NOS: 91  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 90  
LENGTH: 2319  
TYPE: DNA  
ORGANISM: Human  
US-09-058-489-90

Query Match	17.6%	Score 383.2	DB 3	Length 2319
Best Local Similarity	58.5%	Pred. No. 4.6e-101		
Matches 761	Conservative	0	Mismatches 503	Indels 36
Gaps	4			
QY	788	AGACAGCATTAACCTTGACAAATATGACACTTCTTGTGGAAGTGTGACATGATG	847	
DB	529	ACACGGGATTAATTTGAGAAATATGATATATACAGTAGAGCAACGGACGTAACT	588	
QY	848	CACACAGCATTAACCTTGACAAATATGACACTTCTTGTGGAAGTGTGACATGATG	907	
DB	589	GTCCTCAATTAATTTGAGAAATATGATATATACAGTAGAGCAACGGACGTAACT	648	
QY	908	TTCGTAAGTGTGTAATTAACCTTCTTGTGGAAGTGTGACATGATG	967	
DB	649	TTGAACTTAATTTGAGAAATATGATATATACAGTAGAGCAACGGACGTAACT	708	
QY	968	TTGAGAGCATTAACCTTGACAAATATGACACTTCTTGTGGAAGTGTGACATGATG	1027	
DB	709	AGGAGAAAGATTAATTTGAGAAATATGATATATACAGTAGAGCAACGGACGTAACT	768	
QY	1028	TCCATTAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTG	1060	
DB	769	TTTTTACCATTAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCA	828	
QY	1061	GAAATTAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTG	1120	
DB	829	AGGAAATTAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTGAGCAATTTG	888	

QY	1121	TTCCGAGATTGGTCAACACGATTTATTTGGAAACGAGAAATTTTCTTTGGACTGTG	1180
Db	889	CAGGAAATTTGGCTGTACAGATCTATGAGGAAGCCAGAAATTTTCTTACGATCTAGAG	948
QY	1181	TAAGAGCTGTTGTTATATATGGGGGAAACCCAGCTGGGACATTCATTAATTGCACAAATAGTAC	1240
Db	949	TTTCGTCCTTGATGATTTATTTGGTGGTGTGATATTTGGTCACAGATTTGGGACTTAGAAC	1008
QY	1241	AAGCTGTAAATATATTAATGTCCTACTCTCGAAGCTGATGATATCTAGGCAAGAA	1300
Db	1009	GTGGATGCCACTTGTATTAGTACCACTCCAGAGACGCTGTAGGATATGATGAAAGAGAA	1068
QY	1301	AGATTTGGCTCAAAACAGATCAAAATACCTAGTTTGGAGGAAGCTGATGGCATTTGGATA	1360
Db	1069	AGATTGGATTAGACTTCTGCAGAGTACTTAGTGGTGAAGAGCTGATAGATGATGCTGATTA	1128
QY	1361	TGGGTTTTGGTCCAGAAATGGAAGAGTAAATTTCTTGCCCGAGATGGCATCAAGAAC	1420
Db	1129	TGGGATTTGAAACTCAGATACGTCGTATATGTTGAACAAGATATCTATGCCACCAAGAGCG	1188
QY	1421	AGCCGCAAAACCTTATGTTCACTAGTCAACTTTCCAGAGAAATTCAAAGTTGGCTGAG	1480
Db	1189	TTTCGTCACACATGATGTTTATGTCGTACTTTTCCTAAGAAATACAGATGCTTCTCGTG	1248
QY	1481	AGTTTTTAAAGTCAATATATCTGTTTGGTCTGTGGACAAGTGGGTGGACATGTAGAG	1540
Db	1249	ACTTTTGGATGAA--TATATCTTTTGGCTGTAGGACAGATAGGCTCTACTCTTGAGA	1305
QY	1541	ATGTTCAAGACAGACCGTTCTCCAGATGGGCCAGTTCTCAAAAAGAGAAAGCTGTTGAA	1600
Db	1306	ACATCACAACAGAAATAGTTTGGGTGGGAAGACTTATGAATTAACGGTCAATTTCTACTGACA	1355
QY	1601	TTTCGCGAAACATVAGG--GATGAAAGACTATGCTTTGTGAAACTAAGAAAAAG	1657
Db	1366	TTTTAGGTGCAACAGGGAGTATTCATCTTATAGTTTGTGGAGCAAAAAGGGAG	1425
QY	1658	CAGATTTTACTGCAACTTTTCTTTGTCAAGAAAAATTCACATATCAATATCCATGCTG	1717
Db	1426	CAGATTTCCCTGGAGGATTTCTTATACATGAAGATATGCTTGTATCTATGATGAG	1485
QY	1718	ATCCGGAACAGAGAGAGCGGGAGCAGCTCTTGAGATTTTTCGTTGGAAGTGGCCAG	1777
Db	1486	ACCGGTCAACAGAGATGAGAGAGAGGCCCTTACACAGTTTGGCTCAGAGAAAAAGCCCA	1545
QY	1778	TTCTTTGTGCTACTTCAGTACCTGCGAGAGGGCTGATATTTGAANAATGTGCAACATGTA	1837
Db	1546	TTCTATGTGCTAACAGCTGTGGCAGGACGAGAGCATTTCAATTAATGTGACATGTTA	1605
QY	1838	TCAATTTTGCATCTTCTTCTTACATTTGATATGATATGTTCAATCAGATTTGGGCTACTGGTC	1897
Db	1606	TCAATTTTGAATTTGGCAAGATATATTTGAAGAAATATGTCATCTGATTTGGCCGTACAGAC	1665
QY	1898	GTTGAGGAATACTGGCAGAGCAATTTCTTTTGTGATCTTGAATCGGATACCAATTAG	1957
Db	1666	GTATAGGAACCTGGGCTTGGCCACTATTTCTTTAA--TGAATAAATATGATATATA	1722
QY	1958	CACAGCCTCTAGTAAAGATATGACAGATGCTCAACAGATGTTCTTGCATGTTGGAG	2017
Db	1723	CAAGAGATTTGTGATCTTCTGTTAGAACTTAAACAAGAAAGTGCTTCTTGTTGGAAA	1782
QY	2018	AAATTTGCTTATATATATACATCTCTGCTTCACTGTTAG	2057
Db	1783	ATATGGCTTATGAACACCATTCACAAAGGTGGACAGTCCGTGG	1822

RESULT 2  
US-09-058-489-17

; Sequence 17, Application US/09058489

; Patent No. 6103886

; GENERAL INFORMATION:

APPLICANT: Whitehead Institute for Biomedical Research

APPLICANT: Lahn, Bruce

10

▶

```

; APPLICANT: Page, David
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
; TITLE OF INVENTION: the Y Chromosome
; FILE REFERENCE: WHI97-08DA
; CURRENT APPLICATION NUMBER: US/09/058,489
; CURRENT FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/041,877
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 4416
; TYPE: DNA
; ORGANISM: Human
US-09-058-489-17

Query Match      17.6%; Score 383.2; DB 3; Length 4416;
Best Local Similarity 58.5%; Pred. No. 6.5e-10;
Matches 761; Conservative 0; Mismatches 503; Indels 36; Gaps 4

```

QY	788	AGACAGCATTAATCTTGACAACAATACACACTATTCTTTGTGGAGTGCTGCACATGATG	847
Db	529	ACAGGGGATTAATCTTTAGAGAAATATATGATATATACCAGTAGAGCAACCGCAGTA	588
QY	848	CACCACCAAGCAATTTCTGACTTTTGAAGAACTAATCTCTGCAGACACTGAATTAACA	907
Db	589	GTCTCCACATATTTGAATAATTTTAGCATATTTGACATGGGAAATTAATCATGGGGAA	648
QY	908	TTTGCTAAGCTGGTTATATACAGTTACTCTCTGACAAAAATACAGTATTCCTATCATAC	967
Db	649	TTGAACCTTACTCGCTAATACCTGCTCTACTCCAGTCAAAAACATGCCATTCCTATTATTA	708
QY	968	TTGCAGACGAGATTTTATGATGGCTTGCTCTCAAAACAGGCTCGGGAAGA	1022
Db	709	AGGAAAAAAGACATTATAGTGCCTTGCTCCCAACAGATCTGGGAAAACTGCAGATTTTC	768
QY	1028	TTCTACCAATTTTGGCTCA-----TATGATCATGATG	1066
Db	769	TTTATCCCATCTAGTCAGATATATACAGATGGTCCAGAGAACTTTGAAGGCTGTGA	828
QY	1061	GAATTAATCGCAGTCGTTTTAAAGATTGACGAACCAAGTGTATTAATTGTAGACCA	1122
Db	829	AGGAAAAATGAAAGGTATAGGGCGCCGCAAAACAATATCCAAATCTTGTTTTAGCCCA	888
QY	1121	CTCAGAAATTTGGTCAACCCAGATTATTTTGGAGCCAGAAAAATTTTCTTTGGCACTTGTG	1180
Db	889	CAAGAGATTTGGCTGTACAGATCTATAGAGAACCAAAAAATTTTCTTACCGATTTAAG	948
QY	1181	TAAAGCTGTTGTTATATATATGGGGAAACCCAGCTGGACATTCATTCGACAAATAGTAC	1240
Db	949	TTCTCCTTGTGTAGTTATATATGATGGTGTGATATTTGTCACAGATTTGGGACTTAGAAC	1008
QY	1241	AAGCTGTAAATATATTTATGTGCTACTCTCGAAGACATGATGGATATATCAATAGCAAA	1300
Db	1009	GTGATATCCACTTGTATAGTACCCACTCAGGACGCTATAGATGATATGATGAAAGAGAA	1066
QY	1301	AGATTGTCTCAAAACAGATCAAAATACTTATGATTGATGTAAGCTGATGCACTGTTGGATA	1360
Db	1069	AGATTGGATTAGACTTCTCGAAGTACTTATAGTGTGATGTAAGCTGATAGATGCTGAGTA	1122
QY	1361	TGGGTTTTGTCGCAAAATGAAGATTAATTTCTTGCCCGAATGCCATCAAGAAC	1422
Db	1129	TGGGATTTGAAACCTCAGATACGTCGTATAGTTGAACAAGATACATGCGCACCAAAAGGCGC	1188
QY	1421	AGGCCCAAAACCTTATATGTTCAGTGCCAATTTCCAGAGAAATTCMAAAGTTTGGCTGAG	1488
Db	1189	TTGTCTCACACATGATGTTTAGTCTACTTTTCTTAAGGAAATACAGATGCTTGCTGTG	1244
QY	1481	AGTTTTTAAAGTCAAAATATCTGTTGTGCTGTTGGAACAAGTGGTGGACATGTATAG	1544
Db	1249	ACTTTTGGATGAA---TATATCTTTTGGCTGTAGGCAAGTATAGCTCTACCTCTGAGA	1300
QY	1541	ATGTTTACAGACCGTTCTCCAAAGTTGGCCAGTTCTCAAAAAGAAAGCTGTTGAA	1600

Db 1306 ACATCACACAGAAAGTGGTGGAGAGCTTAGATTAACGGTCACTTTCTACTGACA 1365  
Qy 1601 TTCTGCAAAACATVAGG---GATGAAGAATATAGTCTTTGTGAAGTAAGAAAAAG 1657  
Db 1366 TTTTAGTGGCAACGAGAGATTCATCTTACTTAGTGTGTGGAGACCAAAAAAGGAG 1425  
Qy 1658 CAGATTTTACTGCACTTTCTTGTGCAAAAAAATATCAACTACAGTAATCCATGAGT 1717  
Db 1426 CAGATTCCTGAGAGATTTCTTATACATGAAGATATGCTTGTACTAGTATCATGAG 1485  
Qy 1718 ATGGGGAACAGAGAGACGGGAGAGACTCTTGAAGATTTTCTGTTGGAAGTCCAG 1777  
Db 1486 ACCGCTACAGAGATCGAGAGAGGCCCTTCAACGATTCGCTCAGAAAAAGCCAA 1545  
Qy 1778 TTCTGTGCTACTCTAGTACTGCTCCAGAGGGCTGATATTGAANAATGTGCAACATGTA 1837  
Db 1546 TTCTAGTGGCTACAGCTGTGGCAGCAGAGACTAGACATTTTCMAATGTGAGACATGTA 1605  
Qy 1838 TCAATTTGATCTTCTTCTACATGATGATAATGTCAATGAGATGGGGTACTGTC 1897  
Db 1606 TCATTTTGAATTTGCCAAGATATTTGAAGATATGTGCAATCGTATTGGCGTACAGAC 1665  
Qy 1898 GTTGTGGAACTCTGGCAGAGCAATTTCTTTTGTGATCTTGAATCGATTAACCATTTAG 1957  
Db 1666 GTGTAGGAAACCTGGGCTTGGCCACCTCATCTTTAA---TGAAMAAAATATGAATATTA 1722  
Qy 1958 CACAGCTCTAGTAAAGATATGACAGATGCTCAACAGAGATGCTCGATGTTGAG 2017  
Db 1723 CAAGAGATTTGTGATCTTCTTGTAGAGCTAAACAGAGATGCTCTTGTGTGGA 1782  
Qy 2018 AATTTGCTTTAGTACATACATTCCTGCTTCACTGATG 2057  
Db 1783 ATATGCTTATGAACACCACTCAAGGTTGCACTGCG 1822

## RESULT 3

US-09-058-489-14  
; Sequence 14, Application US/09058489  
; Patent No. 6103886  
; GENERAL INFORMATION:  
; APPLICANT: Whitehead Institute for Biomedical Research  
; APPLICANT: Lahn, Bruce  
; APPLICANT: Page, David  
; TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of  
; TITLE OF INVENTION: the Y Chromosome  
; FILE REFERENCE: WHI97-08PA  
; CURRENT APPLICATION NUMBER: US/09/058,489  
; CURRENT FILING DATE: 1998-04-10  
; EARLIER APPLICATION NUMBER: 60/041,877  
; EARLIER FILING DATE: 1997-04-11  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 3408  
; TYPE: DNA  
; ORGANISM: Human  
US-09-058-489-14

Query Match 17.0%; Score 369.4; DB 3; Length 3408;  
Best Local Similarity 57.6%; Pred. No. 5,7e-97;  
Matches 758; Conservative 0; Mismatches 521; Indels 36; Gaps 4;

Qy 773 TCTTGCACATATACAGAGGCACTTAACCTTCGCAAAATAGACACTATTCTTTGGAAG 832  
Db 1305 TCTTTTCTGAGGCAACCTGGGATTAATTTTGAAGAAATGAGATGACATTTCACTTAGG 1364  
Qy 833 TGCTGACATGATGACACACCAATTTGACTTTTGAAGAAGTAATCTGTGAGA 892  
Db 1365 CAACAGGCAACAATGCTCTTCAATATTTGAAGTTTCAAGATGTTGAGATGGAGAAA 1424  
Qy 893 CACTGAATACACATTTGCTAAAGCTGTTACTTAAGCTTACTCTGTGCAAAATACA 952

Db 1425 TTAATGGAACATTTAGCTTACTGCTTAATCTGCCCCCACTCCAGTCCAAAAGCATG 1484  
Qy 953 GTATTCCTATCATATCTTGCAGAGACGATTTGATGCTTGTGCTCAAAACAGGCTTGGA 1012  
Db 1485 CTATTCCTATTAACAAGAAAGAAAGAGACTGATGCTTGTGCTGCAAAACAGGCTTGGA 1544  
Qy 1013 AGACTGCGCTTTTCTCTCAACAAATTTGGCTCATATGATGATGAG----- 1061  
Db 1545 AACTGCAAGATTTCTGTTGCCATCTTGAAGTATTTTATGAGATGCTCAGGCGAGG 1604  
Qy 1062 -----AATACTGCCAGTGGTTTAAAGATGGTGCAGGAACCGAGTGA 1105  
Db 1605 CTTTGAAGGCCATGAAAGAAATGAAAGATGATGGCGCCCAACCAATACCTTCTCT 1664  
Qy 1106 TTAATGAGCACCACATCTGAGAAATGGTCAACAGATTTATTTGGAAGCCAGAAATTTT 1165  
Db 1665 TGGTATTTAGCAACCAAGAGAGATTTGGCAGTACAGATCTACGAAGAACCGAATAATTT 1724  
Qy 1166 CTTTGGGACTGTGTAAGAGCTGTGTTATATATATGAGGGAACCCAGCTGGGACATTCAA 1225  
Db 1725 CATACCGATCTAGAGTTCGCTTGGCGTGTATATGTTGCTGCGATATTTGTCAGAGA 1784  
Qy 1226 TTGACAAATAGTACAAAGCTGTATATATATATGCTACTCTCTGGAAGACTATGATA 1285  
Db 1785 TTGAGACTTGGAACTGATGATGCTATTTGTAGTACGACCTCCAGAGCGTCTAGTATA 1844  
Qy 1286 TCAATGCAAGAAAGATTTGCTCTCAACAGATCAATATCTTATGTTGGATGAGACTG 1345  
Db 1845 TGATGAAAGAGAAAGATTTGATAGACTTTTGCAAAATCTTGCTGTATGATTAACCTG 1904  
Qy 1346 ATGCAATGTTGATATGAGTTTGTGTCAGAAATGAAAGATTAATTTCTTGGCCAGAA 1405  
Db 1905 ATGCAATGTTGATATGAGTTTGTGTCAGAAATGAAAGATTAATTTCTTGGCCAGAA 1464  
Qy 1406 TGGCATTAAGAGAAACAGCGCCAAACCTTATGTTCACTGCAACTTTTCCAGAGAAATTC 1465  
Db 1965 TGCCCTCAAAAGGCTGTCGCACTATGATGTTTGTAGTACTTTTCTTAAGAAATATC 2024  
Qy 1466 AAGGTTGGCTGCGAGATTTTAAAGCAATATCTGTTGTGCTGTTGGAACAAGTG 1525  
Db 2025 AGATGCTGCTCGATTTTCTTATGATGA---TATATCTTCTTGCTGTGGAAGATTTG 2081  
Qy 1526 GTGAGCATGTAGAGATTTCAAGACACCGCTTCCAAAGTTGGCCAGTTCTCAAAAAG 1585  
Db 2082 GCTTCACTCTGAAAACATCACAGAAAGTATGTTGGTGAAGATCACAAACGGT 2141  
Qy 1586 AAAAGCTGTTGAAATTCGCGAAACATAGG---GATGAAGAACTATGCTTTGTG 1642  
Db 2142 CATTTCTGCTTGAACCTCTTAATGCAACAGCAAGATTCATGACCTTATGTTGTG 2201  
Qy 1643 AACTAGAAAAGAGATTTTACTGCACTTTCTTGTGCAAGAAAATATCACTA 1702  
Db 2202 AGACCAAAAGGCTGCAATTTCTGTGAGATTTCTTATACATGAAGATACCAATGTA 2261  
Qy 1703 CAAGTATCATGATGATCGGGAACAGAGAGCGGGAAGCAAGCTTTGAGATTTTGCT 1762  
Db 2262 CAGCATCATGAGAGACGTTTCTCAGAGGATGAGAAAGAGGCCCTTCAACAGTTCCGCT 2321  
Qy 1763 TTGGAAGTCCCACTTCTTGTGCTACTTCACTGAGTCTCCAGAGGCTGATTTGAAA 1822  
Db 2322 CAGGAAAAGAGCCCAATTTTGTGCTACACAGATGAGCAAGAGAGCTGCACTTTCAA 2381  
Qy 1823 ATGTCAGAAAGTATTAATTTGATCTTCTTACATTTGATGATGATGATGATGATA 1882  
Db 2382 ATGTCAGAAAGTATTAATTTGATCTTCTTACATTTGATGATGATGATGATGATA 2441  
Qy 1883 TTGGCGTACTGCTGTGTTGGAATATCTGCAAGCAATTTCTTTTGTGATCTGAAT 1942  
Db 2442 TTGGCTTACGGGACGTTGAGAAACCTTGCGGCAACCTATCTTTAA---CGAGA 2498  
Qy 1943 CGGATTAACATTTAGCACAGCTCTTATGAAGATTTGACAGATGCTCAACAGATGTTG 2002  
Db 2499 GGAACATTAATTAATTAAGATTTGTTGATCTTCTTGTGAAGCTTAACCAAGAGTGC 2558

Qy 2003 CTGCATGCTTGGAGAAATGCCCTTATGATACATACATTCCTGCTTACTGTGTAG 2057  
Db 2559 CGCTTGGTTAGAAACATGGCTTATGAACACACATCAAGGCTAGCAGTCCGTGG 2613

RESULT 4

```

Sequence 3: Application US/09058469
Patent No. 6103886
GENERAL INFORMATION:
APPLICANT: Whitehead Institute for Biomedical Research
APPLICANT: Iahn, Bruce
APPLICANT: Page, David
TITLE OF INVENTION: Genes in the No. 6103886-Recombining Region of
TITLE OF INVENTION: the Y Chromosome
FILE REFERENCE: WHI97-08PA
CURRENT APPLICATION NUMBER: US/09/058,469
CURRENT FILING DATE: 1998-04-10
EARLIER APPLICATION NUMBER: 60/041,877
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 91
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13
LENGTH: 5322
TYPE: DNA
ORGANISM: Human
US-09-058-469-13

```

Query Match	17.0%;	Score 369.4;	DB 3;	Length 5322;
Best Local Similarity	57.6%;	Pred. NO. 7.2e-97;		
Matches 758; Conservative	0;	Mismatches 521;	Indels 36;	Gaps 4

OY	773	TCCTTGGACATTTATGACGACGAGCAATTAACCTTCGACAAATTAAGACACACTTTCTTGCGAAG	83.2
Db	1305	TCCTTTCTCGGAGGCAACACTGGGATTTAATTTTGAGAAATACAGATGACATTCCTCAGTTGAGG	136.4
OY	833	TGTCGACATGATGACACACACAGCAATTCGACTTTTGAAGAACTAATCTCTGTGACA	89.2
Db	1365	CAACGGCAACACACTGCTCTCCACATATTTGAAGTTTCAGTGATGTTGATGGGAGAA	142.4
OY	893	CACGATATACACACTTGTCTAAGCTGGTTATCTAAGCTTAATCTCTGCGAAAATAC	95.2
Db	1425	TTATCATGGGAAACATTTAGCTTACTCCTTACTCGCCCACTCCATGCGAAAAGCATG	148.4
OY	953	GNATTCCTATCACTTCCAGGACGAGATTGATGGCTTGCGTCAACAGGGTCTGGGA	101.2
Db	1485	CTATTTCTTATATCAAGAGAAAAGAAAGACTGATGGCTTGCGCCAAACAGGTTCTGAA	154.4
OY	1013	AGACTGCGGCTTTTCTCCTACCAATTTTGGCTCATATGATGATGATGG-----	106.1
Db	1545	AAACTGACGATTTCTGTGGCCACTTTGAGTCAATTTATTCAGATGTCGACGGCAGG	160.4
OY	1062	-----AATTACTGCCACTCGTTTTAAAGATTGACGAGAACACAGTCTA	110.5
Db	1605	CTTTGAGGGCATGAGAGAAATGAGAAAGTATGGCGCGCGCAACAAATACCAATCTCT	166.4
OY	1106	TTATTTGTAACCAACCTGAGAGATTGTGTCAACCAATTTATTTGGAACCGCAAAATTT	116.5
Db	1665	TGTAATTTAGCACCAACGAGAGGTTGGAGATCAATCTACGAAAGCCACGAAAATTTT	172.4
OY	1166	CTTTTGGGACTTGTGTAAAGCTGTGTGTATATATGAGGGGAAACCCAGCTGGGACATTCAA	122.5
Db	1725	CATACCGATCTAGAGTTGCTCTTGGCGGTTATATGATGGTGGCATATTTGGTCAGCAGA	178.4
OY	1226	TTCGACAAATAGTACAGGCTGTAAATATATATATGTGTACTCTCTGGAAGACTGATGATA	128.5
Db	1785	TTTCGAGACTTGGAAAGTGGATGCCATTTTGTATAGTACCACTCCAGGAAGCTCAATGGGATA	184.4
OY	1286	TCATAGGCAAAAGAAAGATGGTCTCAAAACAGATCAAAATCTAATTTTGGATGGAAGCTG	134.5
Db	1845	TGATGAAGAAGAAAGATGGATTAAGACTTTTGCAAAATCTGTTGTATGATGAAGCTG	190.4

OY	1346	ATCGCATGTTGGAAATATGAGTTGGTTGGTCACAGAAATGAAGAAGTAATTTCTTGCCACGAAA	1405
Db	1905	ATCGAGTTGGATATGATGGGTTTGAGCTCAGATTCCTAGATATGCTGACACAAATATCTA	1964
OY	1406	TGCCATCAAAAGAAACAGCCCAAAACCTTATGTTCAAGTGCACATTTTCCAGAGGAAATTC	1465
Db	1965	TGCCTCCAAAAGGGGTGTCGCCACACTATGATGTTTATGTCTACTTTTCTTAAAGGAAATAC	2024
OY	1466	AAAGTTGGCGTCGACAGATTTTAAAGTCAATATCTGTTTGTGCTGTGGACAAAGTG	1525
Db	2025	AGATGCTGGCTCGATATTTCTTAAATGAA--TATATCTTCTGGCTGTAGGAAGAGTTG	2081
OY	1526	GTGAGCATATGAGATGTTCAAGCAGACCCGTTCCAAAGTTGGCCAGTTCTCAAAAAGAG	1585
Db	2082	GCTCTACCTCGAAAACATCACAGAAAGATGTTGGGTGGAAGAAATCAGACAAACGGT	2141
OY	1586	AAAAGCTCGTTGAATTTCTGGGAAACATAG--GATGAAAGAACTATAGTCTTTGTG	1642
Db	2142	CATTTCTGCTTGACCTCTTAATATGCAACAGCAAGATCACTGACCTTATGTTGTGG	2201
OY	1643	AAACTAAGAAAAAGACAGATTTTATCTGCAATTTTCTTTCGAAAAAAATATCAACTA	1702
Db	2202	AGACCAAAAAGGGTGACAGATTTCTGGAGATTTCTTATACCATGAAGATATGCCATGTA	2261
OY	1703	CAAGTATTCATGTTGATTCGGGAAACAGAGACAGCGGGAGCAAGCTTTGGAGATTTTGGCT	1762
Db	2262	CCACATTCACATGAGAGACCGTTCTAGAGGGATAGAGAAAGAGCCCTTACACAGATTCCGCT	2321
OY	1763	TTGGAAGATGCCCAAGTCTTGTGCTACTTCAGTAGCTGCACAGAGGGCTGGATATTGAAA	1822
Db	2322	CAGAAAAAGGCCAATTTTATGCTGCTACAGACAGTAGACAGCAAGAGACTGGACATTTTCAA	2381
OY	1823	ATGTGCAACATGTTATCAATTTTGAATCTTCTTACCATTTGATGAAATATGTTATGCAA	1882
Db	2382	ATGTGAAACATGTTATCAATTTTATCACTTGCCAAAGTATATGAAAGAAATATGTACATGTA	2441
OY	1883	TTGGCGGACTAGTCGTTGTTGGGAATACTGSCAAGACAAATTTCTTTTTCATCTTGAAAT	1942
Db	2442	TTGGTCTGACGGGACGTGTAGAAAACTTGGCCTGGCAACTCATTTTAA---CGAGA	2498
OY	1943	CGGATTAACCATTTAGCACAGCCTTAGTAAAGTATGACAGATGCTCAACAGAGATGTTT	2002
Db	2499	GGAACATTAATATTAATCAATAGAGATTTGTGATCTTTCTTGTGAAGCTTAACAAAGAGTC	2558
OY	2003	CTGCATGTTGGAGAAATGGCCTTTAGTACATCAATTTCCGCGCTCAGTGGTGG	2057
Db	2559	CGTCTGGTTGAAGAAACATGGCTTATAGAACCCACTACAGAGGTATGACGTGTGG	2613

## RESULT

```

US-09-183-706-42
: Sequence 42, Application US//09183706
: Patent No. 6245525
: GENERAL INFORMATION:
: APPLICANT: Martelange, Valrie
: APPLICANT: De Smet, Charles
: APPLICANT: Boon-Falleur, Thierry
: TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR
: FILE REFERENCE: 10461/77054
: CURRENT APPLICATION NUMBER: US//09/183,706
: CURRENT FILING DATE: 1998-10-30
: EARLIER APPLICATION NUMBER: 09//122,989
: EARLIER FILING DATE: 1999-07-27
: NUMBER OF SEQ ID NOS: 43
: SEQ ID NO 42
: LENGTH: 2365
: TYPE: DNA
: ORGANISM: H. sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (208)...(2151)
US-09-183-706-42

```



Query Match 7.3%; Score 158.2; DB 4; Length 2365;  
 Best Local Similarity 50.1%; Pred. No. 8e-36;  
 Matches 518; Conservative 0; Mismatches 488; Indels 27; Gaps 4;

QY 902 ACAACATGCTGAAGCTGTTATCTAAGCTTACTCCTGCGCAAAATATACAGATTCTTA 961  
 DB 971 AAAACATTTAAAAAGCAGGTTTTCATAAGCCACCTATTCTAGTCACAGCATGGCCCA 1030  
 QY 962 TCATCTTGCAGACGAGATTTGATGCTTGCTCAACAGAGGCTGGGAAGACTGCGG 1021  
 DB 1031 TTGTGTTGCAAGAAATGATCTTATAGAGTAGCCCAAGCTGGAAACGAAAGACATGTT 1090  
 QY 1022 CTTTCTCTTACCAATTTGGCTCAATGATGATGATGATGATGATGATGATGATGAT 1081  
 DB 1091 GTTATTAAATGCTGAGTTTATTCATCTGCTCTTCAAC-----CCAGCTTAAAG 1141  
 QY 1082 AAGAGTTGACGAGAACCGAGTGTATTTGTAGACCAACTCCGAATTTGGTCAACCA 1141  
 DB 1142 GTCAAGGAATGACCCGCGAGTGTAGTCTTAATCTCCACTCCGGAATTGACACTTCA 1201  
 QY 1142 TTTATTGGAAGCCAGAAATTTCTTTGGGACTTGTGTAAGAGCTGTTGTTATATATG 1201  
 DB 1202 TAGAAGAGAAATGTTGCAATATTCATAT---AAGGGCTTGGAGTGTGTTGTTATATG 1258  
 QY 1202 GGGGAACCCAGCTGGGACATTCATTCAGCAATATGTAAGAGCTGTAATATATATATG 1261  
 DB 1259 GTGGTGAATATAGATGACAAATAGAGAGCTTAAAGAGGTGATGATATCATATATG 1318  
 QY 1262 CTACTCTGGAAGACTGATGATATCTATAGGCAAAAGAAAGTTGGTCTCAACAGATCA 1321  
 DB 1319 CAATCCCGGAAGATGAAATGATCTGCAATGAGTAACTTGTCAATCTCAAGAAATTA 1378  
 QY 1322 AATACTAGTTTGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381  
 DB 1379 CCACTTGGTTTATGATGAGAGCAAGATGTTGGATGATGATGATGATGATGATGAT 1438  
 QY 1382 AGAAGTTAATTTCTTGCCAGGAATGCCATCAAGGAACAGCGCAACCTTATATGTTCA 1441  
 DB 1439 TGAAGATTTTGT-----TAGATGGGCCAGATAGGCAACAGTTATATGCCA 1486  
 QY 1442 GTGCACTTTTCCAGAGAAATTCAAAGTTGGCTGCAAGTTT---AAAGTCAAAAT 1498  
 DB 1487 GTGCTACATGGCTCATTCATTCATGCTGCTGCAATCTTATTTAAGAAACCAATGA 1546  
 QY 1499 ATCTGTTGTTGCTGTTGGAACAAGTGGTGAGCATATAGAGATTCACAGACCGTTC 1558  
 DB 1547 TTGTCTATGTTGTTGATCTGATGATCTAGTGTGATGATGATGATGATGATGATGAT 1606  
 QY 1559 TCCAGTTGGCCAGTTCTCAAAAAGAGAAAGCTGTTGAATTTCTGCAAAACATAGGGG 1618  
 DB 1607 TAACCAACGAGGAAGAGAAATGAGATGCAATGCAAACTTTTCTACAGATATGTCACA 1666  
 QY 1619 ATGAAAGACTATGCTTTTGTGTAACCTAAGAAAAAGCAATTTTACTGCAACTTTTC 1678  
 DB 1667 CAGACAAAGCATATGCTTCTGTTCTCGAAAAGCTGTTGGGATCACCTTATCAAGTACC 1726  
 QY 1679 TTGTCAAGAAAAATATCAACTACAGTATCCATGATGATGATGATGATGATGATGATGAT 1738  
 DB 1727 TAATCTTGAATATATACGTAGAGTCTCTGCAAGAGATGAGAAACAGAGAGATGGGG 1786  
 QY 1739 AGCAAGCTCTTGAGATTTTGCCTTTGGAAGTCCAGATTTCTGTTGTTACTTCACTAG 1798  
 DB 1787 AGAAGCATTTAGAGAACTTTAAAAAGGCAAGGTAAGTAATCTATTTGCAACTGATAG 1846  
 QY 1799 CTGCGAGAGGCTGATATGAAATGTCAACATGTTATCATTTTGTATCTTCTTCTA 1858  
 DB 1847 CCTCTAAGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1906  
 QY 1859 CCATTTGATATATGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1918  
 DB 1907 ATATTGAAGAAATGCTACACCGAATAGGCGCAGGGAAGAGCAGGAGAGATGCTGTTT 1966  
 QY 1919 CAATTTCTCTTTT 1931

DB 1967 CCATTTCAACTTT 1979

RESULT 6  
 US-09-567-995-42  
 ; Sequence 42, Application US/09567995  
 ; Patent No. 6303756  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Martelange, Val,rie  
 ; APPLICANT: De Smet, Charles  
 ; APPLICANT: Boon-Falleur, Thierry  
 ; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
 ; FILE REFERENCE: L0461/7054  
 ; CURRENT APPLICATION NUMBER: US/09/567,995  
 ; CURRENT FILING DATE: 2000-05-10  
 ; PRIOR APPLICATION NUMBER: 09/183,706  
 ; PRIOR FILING DATE: 1998-10-30  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SEQ ID NO 42  
 ; LENGTH: 2365  
 ; TYPE: DNA  
 ; ORGANISM: H. sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (208)...(2151)  
 US-09-567-995-42

Query Match 7.3%; Score 158.2; DB 4; Length 2365;  
 Best Local Similarity 50.1%; Pred. No. 8e-36;  
 Matches 518; Conservative 0; Mismatches 488; Indels 27; Gaps 4;

QY 902 ACAACATGCTGAAGCTGTTATCTAAGCTTACTCCTGCGCAAAATATACAGATTCTTA 961  
 DB 971 AAAACATTTAAAAAGCAGGTTTTCATAAGCCACCTATTCTAGTCACAGCATGGCCCA 1030  
 QY 962 TCATCTTGCAGACGAGATTTGATGCTTGCTCAACAGAGGCTGGGAAGACTGCGG 1021  
 DB 1031 TTGTGTTGCAAGAAATGATCTTATAGAGTAGCCCAAGCTGGAAACGAAAGACATGTT 1090  
 QY 1022 CTTTCTCTTACCAATTTGGCTCAATGATGATGATGATGATGATGATGATGATGATGAT 1081  
 DB 1091 GTTATTAAATGCTGAGTTTATTCATCTGCTCTTCAAC-----CCAGCTTAAAG 1141  
 QY 1082 AAGAGTTGACGAGAACCGAGTGTATTTGTAGACCAACTCCGAATTTGGTCAACCA 1141  
 DB 1142 GTCAAGGAATGACCCGCGAGTGTAGTCTTAATCTCCACTCCGGAATTGACACTTCA 1201  
 QY 1142 TTTATTGGAAGCCAGAAATTTCTTTGGGACTTGTGTAAGAGCTGTTGTTATATATG 1201  
 DB 1202 TAGAAGAGAAATGTTGCAATATTCATAT---AAGGGCTTGGAGTGTGTTGTTATATG 1258  
 QY 1202 GGGGAACCCAGCTGGGACATTCATTCAGCAATATGTAAGAGCTGTAATATATATATG 1261  
 DB 1259 GTGGTGAATATAGATGACAAATAGAGAGCTTAAAGAGGTGATGATATCATATATG 1318  
 QY 1262 CTACTCTGGAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1321  
 DB 1319 CAATCCCGGAAGATGAAATGATCTGCAATGAGTAACTTGTCAATCTCAAGAAATTA 1378  
 QY 1322 AATACTAGTTTGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381  
 DB 1379 CCACTTGGTTTATGATGAGAGCAAGATGTTGGAATGATGATGATGATGATGATGAT 1438  
 QY 1382 AGAAGTTAATTTCTTGCCAGGAATGCCATCAAGGAACAGCGCAACCTTATATGTTCA 1441  
 DB 1439 TGAAGATTTTGT-----TAGATGGGCCAGATAGGCAACAGTTATATGCCA 1486  
 QY 1442 GTGCACTTTTCCAGAGAAATTCAAAGTTGGCTGCAAGTTT---AAAGTCAAAAT 1498  
 DB 1487 GTGCTACATGGCTCATTCATTCATGCTGCTGCAATCTTATTTAAGAAACCAATGA 1546  
 QY 1499 ATCTGTTGTTGCTGTTGGAACAAGTGGTGAGCATATGATGATGATGATGATGATGATGAT 1558

```

Db      1547 TTGCTATGTTGACATGATGATGATGCTGTAAGTCACTGACGAAATAATATTG 1606
QY      1559 TCCAGTTGGCCAGTTCTCAAAAAGAGAAAAGCTGTTGAATTCTGCGAAACATAGGG 1618
Db      1607 TAACACCGAGGAGAGAAAAGAGTCAACAGCAACTTTCTCAGAGTATGTCATCA 1666
QY      1619 ATGAAAGAACTATGCTGTTGTTGAACTAGAAAAGAGAGATTTTACGCACTTTTC 1678
Db      1667 CAGCAAAAGTCATTTGCTCTTCTGTTCTGAAAAGCTGTTGCGATCACTTATCAAGTACC 1726
QY      1679 TTGTCAAGAAAAAATATCAACTCAAGATTCATGTTGATCGGAAAGAGAGAGCGGG 1738
Db      1727 TAATACCTGGAATATATACAGTAGAGTCTCTGATGAGATAGAGAAAGAGAGATCGGG 1786
QY      1739 AGCAAGCTCTTGAGATTTTGGCTTTGGAAGTCCAGTTCTGTTGCTACTTCACTAG 1798
Db      1787 AGAAAGCATTAAGAACCTTTTAAACAGGCAAGTAGAGAAATTAATTGCAACTGATCTAG 1846
QY      1799 CTGCGAGAGGCTGATATTAATGAATGCAACATGTTATCAATTTTGTATCTTCTCTA 1858
Db      1847 CCTTAAGAGACTGATGTCATACATGCTTACACATGCTATTAATTTGACTTTCCAGGA 1906
QY      1859 CCATTGATGATATGTTATGAAATGAGCGCTACTGCTGTTGAGGAATACCTGCAAG 1918
Db      1907 ATATTGAAGATAGTACACCGAATAGGCGGCAAGGAGAGAGAGAGAGCTGTGTTT 1966
QY      1919 CAATTTCTTTT 1931
Db      1967 CCATTACAACTTT 1979

```

## RESULT 7

```

PCT-US96-05320A-894
; Sequence 894, Application PC/TUS9605320A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences
; APPLICANT: 9410 Key West Avenue
; APPLICANT: Rockville, MD 20850
; APPLICANT: United States of America
; APPLICANT: Johns Hopkins University
; APPLICANT: 720 Rutland Avenue
; APPLICANT: Baltimore, MD 21205
; APPLICANT: United States of America
; APPLICANT: Mark D. Adams
; APPLICANT: Owen White
; APPLICANT: Hamilton O. Smith
; APPLICANT: J. Craig Venter
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome,
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20003-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05320A
; FILING DATE: April 12, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Eric K. Steffe
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.014PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 894:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US96-05320A-894

Query Match      5.7%; Score 123.6; DB 5; Length 1254;
Best Local Similarity 47.1%; Pred. No. 6.1e-26;
Matches 495; Conservative 0; Mismatches 539; Indels 18; Gaps 3;

QY      908 TTGCTAAGCTGTTATCTACTAAGCTTACTCTGTCGAAAATACAGTATTCCTATCATAC 967
Db      80 TGGCAAAAAGAGCTTGAATTTTGTACCCCAATTCAGGCTTATCTGCTATCAGTT 139
QY      968 TTGAGAGAGAGATTTGATGCTGTTGCTTAAACAGGCTTGGGAAAGACTGCGCTTTTC 1027
Db      140 TAAATGACGAGATGTCGACAGCAAGCTCAACTGTTACAGGCAAGACATGCTTTT 199
QY      1028 TCTTACCAATTTGGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
Db      200 TAAAGGCTACTTTTACCATCTTTTAACTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
QY      1088 TGCAGAGACAGAGATGATATTTGATGACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1147
Db      251 ATCTCAACCAAGAGCTTATTTTGAACCTACTGAGAGATTAAGGAGTACAGATTAGTA 310
QY      1148 TGAAGCCAGAAATTTTCTTTGGAGCTTGTGTAAGAGCTGTTTATATATGAGGGA 1207
Db      311 ATACGAGAAATTTCTTGCAAAAGCGAGTGAATTAAGACCGACCTGCTATGAGGCG 370
QY      1208 CCGAGCTGGACATTAATTCGACAAATGTAAGGCTGTAATATATATGCTACTC 1267
Db      371 ATGTTATGATTAACAACATCAAGCGATGAGCTGCGATATTTGATGAGTACGA 430
QY      1268 CTGGAAGACTGATGATATCATATAGGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1327
Db      431 CGGGGAGATCATTTGATATGTAAGCAAGCGTAATGTTTGAATGAATCAAGTTG 490
QY      1328 TAGTTTGAATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387
Db      491 TCGGTGATGATGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550
QY      1388 TAAATTTCTCCAGAGATGCGCATCAAGAGACAGGCGCAACCCCTATGTTGAGTCA 1447
Db      551 TATTCGTA-----ATGCCCGCTCCGCAAGCTCTTTAAGATGTTATTTTCACGA 604
QY      1448 CTTTCCAGAGAAATCAAGGTTGCTGAGATTTTAAAGTCAATATCTGTTG 1507
Db      605 CGCTTCTTAAAGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
QY      1508 TTGCTGTTGACAAAGTGGTGAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1567
Db      662 TTGAATTTGAACAGAAACAAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
QY      1568 GCGAGTCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1627
Db      722 CTAATCAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
QY      1628 CTAAGCTTTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1687
Db      782 GTATGATTTTGGAAATACAGAAACATGCTGTAAGAAATTTGGGAGTATTTGGCGGCTG 841
QY      1688 AAAAATATCAACTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1747
Db      842 ATGGGATCGTGTGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901

```

QY	1748	TTGGAGATTTTCGCTTTGGAAAGCCAGTCTTGTTGCTACTGACAGTCCAG	1807
Db	902	TAAACAATTACTGATGCTGATTTGGATATTTTAGTGGCAACAGATGTGCTCTGTG	961
QY	1808	GGCTGATATTGAAATGTGCACATGTTATCAATTTTGATCTCTTCAACATTATG	1867
Db	962	GCTTGCAATTTTCTGATGTGACGCATGTTTCATTATTAATTAACCGAGATGAGAAAG	1021
QY	1868	AATATGTTTCATCCGAATGGGGCTACTGTGCTTGTGGGAATATCTGCAGACAAATTTCTT	1927
Db	1022	ATTATGTTCACCGAATTTGGGGCTACTGCAGACAGGGGAAAGTGTGTTTGATTTAGTT	1081
QY	1928	TTTTGATCTGAATCGGATPACCAATTTAGCA	1959
Db	1082	TTCCTGTGGAAGATATATGCCATGAATTTACCA	1113

RESULT 8

US-08-961-527-203/c  
; Sequence 203, Application US/08961527

GENERAL INFORMATION:

APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

FILING DATE: ;

ATTORNEY/AGENT

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PE  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEPHONE: (301) 309-8512  
TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 203:

; SEQUENCE CHARACTERISTICS:

LENGTH: 3230 base pairs

TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear

US-08-961-527-203

C  
C  
C  
C  
C  
P  
C  
P  
-  
P  
C  
L

Query Match	5.48;
-------------	-------

Best Local Simi

Matches 505;

2000

1111 1.68 863 57

Accession	Sequence	Position
Db	TGAAATTTTAAAGCATTTAAACCTGTCTGTCTGATTTTGCTAGACAGAAATTAATAAAAGCTGGTT	1851
Qy	ATACTAAGCTTACTCTCTCTGTGCAAAATATCAGATATTTCTATCATATCTTGACAGACGAGTT	982
Db	TTGTGAAAGCTAGTCTCTATTCACAGAACCAACTATTCCTCTTTTGACCCCTTGAAAGCAGAGATG	1791
Qy	TGATGAGCTTGCTGCTCAAAACAGAGGTCGGGAAGACTGCGGGCTTTTCTCTACCAATTTTGG	1042

Db	1790	TTATCGGTCAAGCTCACAAGTGGTACAGAGAAAATCGAGGCTTTGGGTGGCTACCCCTTG	17311
Oy	1043	CTCATATGATGATGATGAAATACTGCCAGTCTTTTAAAGAGTTGCAGAACACAGACT	11022
Db	1730	AAAAAATCCGTACAGAAAGAACGCACTATCCAAAGCCTTGG-----	16922
Oy	1103	GTATATTGTAGACACCAACTCGAAGATTTGGTCACACAGATTTATTGGAAAGCCAGAAAT	11622
Db	1691	-----TCATCGCTCCAACTGCTGTAAGTCTAGCTGCTCCAAAGTCAAGAAAGACTTTCCGCT	16382
Oy	1163	TTTTCTTTTGGGACTTGTGAAGAGCTTGTATATATATGGGGGAAACCCAGCTGGGACATT	12222
Db	1637	TTGGTGTGTAAGAGGAGTCAAAAGTCCGTTACGATATATGGGGATTCAGACATTGAAAAC	15782
Oy	1223	CAATTGCACAAATATGTACAAAGGCTGTATATATATATGTGCTACTCTCGGAAGACTGATGC	12822
Db	1577	AAATTAAGGCTCTTAAATCTGGTGCCCATATTTGGTGGGAAATCCAGGTGCTCTTGG	15182
Oy	1283	ATATCATAGGCAAAAGAAAGATTTGGTCTCAACACGATCAATATCTTAATTTGGATGAAG	13422
Db	1517	ACTTGATTTAAACGACAGGSCCTTGAAATTTACACAGCATTTGAAACCCCTTATCCTTGACAGAG	14582
Oy	1343	CTGATCGCATGTTGGATATGGGTTTGGTCCAGAAATGAAAGAAATTAATTTCTTGCCAG	14022
Db	1457	CGGATGAATATGCTTAACATGGGCTTCTTGAAAGCATCGAAGCATTATTTCC-----	14062
Oy	1403	GAAATGCCATCAAGAGAACAGGCGCCAAACCTTATGTTCACTGCACACTTTTCCAGAGAA	14622
Db	1405	-----CCGTGTACCTGGAACCCGTCAAACTTTGCTTTTCTCAGCAACATATGCAATGCCA	13502
Oy	1463	TTCAAAAGTTGGGCTGCGACAGATTTTAAAGTCAAAATATCTGTTGTTGCTGTTGGACAAAG	15222
Db	1349	TCAACGCTATCGGTTCTCAGTTTATGAAAGCCCTCGTAAC---ATGTCAAGATTGGCGCTA	12932
Oy	1523	TGGGTGAGCATGTAGAGATGTTTCAGACAGCCGTTCTCCAAAGTTGGGCAAGTTCTCAAAA	15822
Db	1292	AGGAATTTGACAAACAGATTTGGTTGACAGATATATATCCGTGTAAAGAAACAAAGAAAT	12332
Oy	1583	GAGAAAGCTCTTGAAATTTCTGGGAAACATAGGGAGTGAAGAACTATATGTCCTTTGTG	16422
Db	1232	TTGACACCATGACTCGTCTCATGTAGATGGGCACAAACAGAACTCGTATATGTATTTGGTCTC	11732
Oy	1643	AAACTAGAAAAAAGACAGATTTTAACTGCAACTTTTCTTTGTCAGAAAAAATATCAACTA	17022
Db	1172	GTACCAAAAGCCCGTGTGGATGAATAGCTGTGGTTTGAATAATTCGTGGCTTCCGTGACAG	11132
Oy	1703	CAAGTATTCATGTGATCGGGAAACAGAGAGCGGAGCAAGCTTTGGAGATTTTGGCT	17622
Db	1112	AAGGAATTCATGGGCACTTAGACCAAAACAAAGCTTTCGTGTCCTTCGTGTATTAATA	10532
Oy	1763	TTGGAAGTCCCAAGTCTTGTGCTATCTTCAGTAGCTGCCAAGGGCTGGATTTTGAAA	18222
Db	1052	ATGGCAATCTTGATGTTTGGTGGCAGACAGCTTGGAGCGCGGTGTGGATATTTTCAG	9932
Oy	1823	ATGTCGAACATGTTATCAATTTTGATCTTTCCTTACCATTTGATGAAATATGTTATCGAA	18822
Db	992	GTGTGACCACTGTCTAACATCAGATATTTCCACAAGATCCTGAGAGTTATGTTCAACGTA	9332
Oy	1883	TTGGGCGTACTGGTGTGGGGAATCTGGCAGAGCAATTTCTTTTTTGATC	19362
Db	932	TCGGTCTGTACAGGCTGCTGTGTAAAGTCAGGTCAATATATTAATCTTTGTTGCTC	8792

## RESULT 9

US-09-208-742-3  
; Sequence 3, Application US/09208742

## ; GENERAL INFORMATION:

```

; APPLICANT: Kaufmann, Joerg
; TITLE OF INVENTION: C1F150/h7AF1150 is Necessary for Cell
; TITLE OF INVENTION: Cycle Progression
; FILE REFERENCE: 1453.002

```

```

CURRENT APPLICATION NUMBER: US/09/208,742
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 3.0.
SEQ ID NO 3
LENGTH: 3825
TYPE: DNA
ORGANISM: human
US-09-208-742-3
```

Query Match	5.3%	Score 115.2;	DB 4;	Length 3825;
Best Local Similarity	46.3%;	Pred. No. 3e-23;		
Matches 513; Conservative	0;	Mismatches 568;	Indels 27;	Gaps 3,

OY	832	TTGTCGGACATGATGACCAACACACCAATTCCTGACTTTTGAAGAAGCTAATCTCTGAC	891
Db	862	GTCCTGGTCGTCACCTCTTGACACGAGAAGTAGCTTTGCTCATTTTGGTTTACGAA	921
OY	892	AACCTGAATMACACACATTTGTAAGCTGGTTATTAAGCTTACTCTGTGCAAAATATAC	951
Db	922	CAACTTATGACACCAATTCGGAAATCTGAATACACAGGCCACTCCAATATAGTCCAG	981
OY	952	AGTATTCCTATCATCTTGACGAGAGATTTGATGGCTGTGCTCAACACAGGCTCGG	1011
Db	982	GGTGTGCTGTGGCAATTAAGTGTAGAGCAATTTGGATTGTATGCCAAACACAGTAGTGG	1041
OY	1012	AAGACTGGCGCTTTCTCTCAACCAATTTTGGCTCATATGATGCATATGATTAATCTGC	1071
Db	1042	AAAATCGACGCTTCATTTGGCCATGTGTATCATATATATGAGCA-----G	1089
OY	1072	AGTCGTTTTAAAGAGTTGACGAGAACACAGTGTATTTATTTGACACCAATCTGAGATTG	1131
Db	1090	AAGAGTTGGAAACCAAGTATGACCAATTCGACGATTTGTGTCTTACACGAGAGCTT	1141
OY	1132	GTCACACAGATTTATTTGGAAGCCAGAAATTTTCTTTTGGAGCTGTGTATAGACTGTT	1191
Db	1150	TGCCACGAGATTCACAGAAATGTATAGCGTTTGGAAAAAGCATATATCTTCGATCAGTG	1201
OY	1192	GTTATATATGGGGGAACCCAGCTGGGACATTCATTCGACAAATATGTACAGAGCTGTAT	1251
Db	1210	GCCGTATATGGGAGGAGAGTATGTGGGAGCGGCAAGGCCCTTCAGAGGGGGCAGAG	1261
OY	1252	ATATATATGCACTCTGGAAGCTGATGATATCATATAGCAAAAGAAAGTTGTC	1311
Db	1270	ATTGTTGTGTATCCCGACAGCTGATAGTATGATGTAAAGAAAGAAAGCTACCAATCTT	1321
OY	1312	AAACGATCAATATCTTAGTTTGGATGAAGCTGATCGCATGTGGATATGGGTTTGGT	1371
Db	1330	CAAAGAGTCTTATCTTGTTTGTATGAAAGCAATCGAATGTTTGAACATGGGATTTG--	1381
OY	1372	CCAGAAATGAAAGATTAATTTCTTGCCAGGAATGCCATCAAAGAACAGGCCAACC	1431
Db	1388	-----AGTACCAAGTTGCATTCATGGAAGTCATGTTGCTCTGCACAGGAGACT	1431
OY	1432	CTTATGTCAGTCAAACTTTTCCAGAGAAATTCAAAGTTGGCTGCAGAGTTTAAAG	1491
Db	1438	CTCTATTATTAAGTCAACTTTTCCGAAAGAAATTTGAAAGTTGGCCAGAGACATCTGATC	1491
OY	1492	TCAATTAATCTGTTTGTGCTGTGGAACAAGTGGTGACATGTAGAGATTTCAAGCAG	1551
Db	1498	GACCTTATTCAGTGTGTGAGGAGAAATATTTGAGAGGCAAAATGAAGATGTGACACAGATT	1551
OY	1552	ACCGTTCTCCAAAGTTGGCCAGTTCTCAAAAAAGAGAAAGCTGTGGAATTCCTGGGAAC	1611
Db	1558	GTCGAGATTTCTCAATTCGAGACTTATGTAAATGGAATGGCTTACCCGGGCTGTGTAGAA	1611
OY	1612	ATAGGGATGAAAGAACTATGGTCTTTGTTGAACCTAAGAA--AAAGCAGATTTTACT	1661
Db	1618	TTTACCTCTTCAGGAGTGTCTCTCTCTTTGTATCAAAAAAGCCAAATGCTGAAGAGCTA	1671
OY	1669	GCAACTTTTCTTTGTCCAGAAAAAATATCACTACAAGTATCCATGTGTATCGGGAACAG	1721
Db	1678	GCGAATTAACCTTAAACAGAGGGGTATATCTTGGGCTGCTCATAGGGATATGTAGATCG	1731

Qy	1729	GGAGAGGGGGAGCAAGCTCTGGAGATTTTCGCTTTGGAAAGTCCCAAGTCTTGTTGCT	1788
Db	1738	AGTAGAGAAACAAAGGTCATTTTCAGACTTTTAAGAAAAAGCAATCCCAAGTCTTGTTGCC	1797
Qy	1789	ACTTCAGTAGCTGCCAAGGGCTGGATTTGAAATGNCACATGTTATCAATTTTGAT	1848
Db	1798	ACAATATGTGACGCCCGTGGTCTGGACATTCCTTCATTTAAGACTGTCTTAATTAATGAT	1857
Qy	1849	CTTCCCTTCACCAATGATGAAATATGTCATCGAATTGGGCGTACTGTCGTTGGCAAT	1908
Db	1858	GTGCGACGGGCAATTGATACGACACATCACAGGATTTGGCGCACAGGAAGCGCGTGAG	1917
Qy	1909	ACTGGCAGACAATTTCTTTTGTATC	1936
Db	1918	AAAGGTGTGCTTAATCCCTACTACTC	1945

```

RESULT 10
US-09-134-001C-1425
; Sequence 1425, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1425
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1425

```

Query Match	4.8#	Score	10.3	2	DB	4	Length	1567
Best Local Similarity	48.5%	Pred.	5	76	20			
Matches	411	Conservative	0	Mismatches	423	Indels	14	Gaps
Qy	1090	CAGAACCCAGAGTGTATTTATTTGTAGCAACCACTTGAGAAATGTCACCAAGATTTATTTG	1149					
Dp	259	CAGGAGTTCATCATTTATTTCTTGCGCCACGAGAGGCTTGCTATGACAGTGTAGC	318					
Qy	1150	GAAGCCAGAAATTTCTTTTGGGACTGTGTAGAGCTGTGTATATATATGAGGGGAAC	1209					
Dp	319	CAATTTAGAGGCTTCAGTAAGGTCAAAATATCAAGTTGTTACAGCTTTGGTGGTATG	378					
Qy	1210	CAGCTGGACATTCATTTGCAAAATAGTACAGGCTGTAAATATATATATGTCTACTCT	1265					
Dp	379	CCTATTGAAAGACAGATTTAAGACATTTAAAAAGAGGCCCTCAAAATTTGTGTAGTACGCG	438					
Qy	1270	GGAAGACTGATGAT-ATCATAGCCAAAGAAAGATTTGCTTCAACAGATCAATATCTT	1328					
Dp	439	GGAAGAGTTATTTGATCATCTTAAATGCTGTACACTTAAACTCAAGATTTACACGTT	498					
Qy	1329	AGTTTGGATGAACTGATCGCATGTTGGATATGGGTTTGGTCCAGAAATGAAGAACTT	1388					
Dp	499	A-TTTTGAATGATACAGATGAATGAATGATGAATATGGAATTCATGCATGATATGAGATTAT	557					
Qy	1389	AATTTCTTGCCAGAGATGCCATCAAGGAACAGGCCCAAAACCTTATGTTCAGTGCAC	1448					
Dp	558	TAT-----GGATAAATTCAGAGTCGAACCAAGTCMAACTATGTATTTTCTGTAC	608					
Qy	1449	TTTTCCAGAGAAATTTCAAAGTGTGGCTGCAGAGTTTAAAGTCAAAATATCTGTTGT	1508					
Dp	609	GATGCCAAAAGCATTCAGAAATTTGTACACAGATTTTATGAAAGCTCTTAATAATATATTA	668					
Qy	1509	TGCTGTGGACAGTGGGTGAGAGATGTAGAGATGTTTCAGACAGCCGTTCTTCCAGTTGG	1566					



```

Qy 917 CTGGTATACAGCTTACTCTGTGCAAAAATACAGATTCCTATCATCTGCAGAC 976
Db 86 TCAATTTGAAAAACCACTGAAATCCAAATAGATCATCCGAGAACTTTAAAGAA 145
Qy 977 GAGATTGATGCTGTGCTCAACAGGGCTGTGGAAAGCTGCGCTTTCTCTACCA 1036
Db 146 CAAATTTAATAGCAATCTCAAACTGGAAGTCAACGCTTTCTTTTACCAT 205
Qy 1037 TTTTGGCTCATATGATGATGGAATTAATGCGCACTGTTTAAAGATTGAGAAC 1096
Db 206 TAAATCAACTATAGAAAGTACA-----TTCAAGAGC 238
Qy 1097 CAGAGTATATATGATGACCAACTGAGATTTGTCACAGATTTATTTGGAAACCA 1156
Db 239 CACAGCACTGCTAGTACCTCCAACTGCACTGCTGCAACATATACAGTTGCTA 298
Qy 1157 GAAAAATTTCT- -TTTGGACCTTGTGTAGAGCTGTGTATATATGAGGGAACCCAGC 1213
Db 299 TGCATTTAGTAAATTCAAAAAGATTAATGTAACCTTTTCATTGTGTGTCGATT 358
Qy 1214 TGGACATTCATTCGCAATATAGTACAGGCTATATATATATGCTACTCTGGA 1273
Db 359 TAGAAAAAGATTAACCAAGATGAGCCATCAACCAACTATTATGTGTACACCA 418
Qy 1274 GACTGATGATATCATATGCGCAAGAAAGATTTGCTCAACAGATCAATTAATTAGTTT 1333
Db 419 GAATTAATGATTTAGCATTCAGCTTATCTTCATGACATTTAGCGCTATATTA 478
Qy 1334 TGGATGAGCTGATCGCATGTTGGATATGGCTTTTGTCCAGAAATGAGAAATTT 1393
Db 479 TAGATGAAGCTGATTTATGATGATGACCTGCTCATTTGAAGATTTGACCA----- 530
Qy 1394 CTGCGCCAGAAATCCATCAAGAAACAGCCCAACCCCTTATGTCAGGCACTTTTC 1453
Db 531 -TATTTGAGCGAGATTTGATGATGATAAATGCTCATAGCGGATTTAGTCAACAATAC 589
Qy 1454 CAGAGAAATTCAGAGGTGCTGACAGATTTTAAAGTCAAAATTAATCTGTTTGTCTG 1513
Db 590 CTAAATCATTAACAACATTTTAAATTAATTTTAAGTCAACGAATTTGTAGAAATTTG 649
Qy 1514 TTGACAAAGTGGGTGAGCATGTAAGATGTTGACAGACCGTTTCCCAAGTTGGCAGT 1573
Db 650 ATGCAAAAGCTCATATATTAAGAAATATCGAATTTTATCTATCTTACAAAAG--TT 706
Qy 1574 TCTCAAAAAGAAAGCTGCTGAATTCGCAAAACATAGGGAGTGAAGAACTATGG 1633
Db 707 CTGCTAAGTAGATTAACATTTGAAATGATATATGAAATCTTATCTATGATTA 766
Qy 1634 TCTTTGTTGAACTAGAAAGAAAGCAGATTTTACCTGCAACTTTTCTTGTCAAGAA 1693
Db 767 TTTTCTGTAACGTCGTAAGAAATCCGATGATTTGGCAGACCTTTAATTAAGAAAGAA 826
Qy 1694 TATCAACTACAAAGTATCATGCTGATCGGGAACAGAGAGCGGAGCAAGCTTTGAG 1753
Db 827 TTAATAATAGATGATTCATGCTGTTTAAACCAAGGAACGTAAACCAATTAAGAAA 886
Qy 1754 ATTTCCTTTGGAAGTGGCCAGTCTTGTGCTACTTCAGTAGCTGCCAGAGGGCTGG 1813
Db 887 GAATTAAGAAATTTGATTTTCAATTTGTCATTCAGAGGATCTTCTAGAGAAATAG 946
Qy 1814 ATATTGAAATGTCACACATGTTATCAATTTGATCTTCTTACCATGATGATATAG 1873
Db 947 ATATTGAAGGCGTAAGTCATTTTATTTGATGTAATGTAACCAATGATATGATTTCTTCA 1006
Qy 1874 TTCAATGAAATGGCGTATGCTGTGTGGAAATAGTGGCAGACAAATTTCTTTTGTG 1933
Db 1007 CACATCGGTAAGTGAAGAGAGATTAATTAAGAGTGAAGTGAAGTCAATTAATATA 1066
Qy 1934 ATCTTGAATCGATTAACATTTAGCAGGCTCTAGTAAAGTATTTGACAGATCTCAAC 1993
Db 1067 GTCTGATGAAGAAAGTATATTAATCTTATTTGAAGACAGAGGTTATTAATTTGAAATG 1126

```

```

Qy 1994 AGGATGCT 2001
Db 1127 TAGATATT 1134

RESULT 13
US-09-722-575-126
; Sequence 126, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yugui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 126
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Human
US-09-722-575-126

Query Match 4.2%; Score 90.8; DB 4; Length 490;
Best Local Similarity 53.4%; Pred. No. 1,2e-16;
Matches 222; Conservative 0; Mismatches 182; Indels 12; Gaps 1;

Qy 973 GACAGATTTGATGCTGTGCTCAACAGGCTGGGAAGCTGCGCTTTTCTCTTA 1032
Db 34 GATTTGATATGTTGATGAGTGCAGACACTGATCTGGAAAAACATGTTTGTCTT 93
Qy 1033 CCAATTTGGCTCATATGATGATGATGAATTAATCTGCAAGTCTTTTAAAGATTGCA 1092
Db 94 CTGCACTGTCACATCAATCATCAATCATCTCTAGAGAGG-----CGAT 141
Qy 1093 GAACGAGATGATTTTATGAGCAACACTCGAATGCTCAACAGATTTATTTGAA 1152
Db 142 GGGCTATTTGTTGTTGCTGCGCACCACTCGGAACTGGCCCAACAGGTGCGCAAGTA 201
Qy 1153 GCCAGAAATTTCTTTTGGAGCTTGTATAGAGCTGTTGTTATATATATGAGGGAACCCAG 1212
Db 202 GCTGCTGAATTTGTAAGAGATGCTGCTGAAGTCTACTGTTATCTACGCTGTGCTCT 261
Qy 1213 CTGGAACATTCATTCGACAAATATGTAACAAGCTGTATATATATATGCTACTCTGGA 1272
Db 262 AAGGACCAACAAATATGATGTTGAGAGAGGTTGGAATCTGTATTCGAACACCTGGA 321
Qy 1273 AGCTGATGATTCATAGCAAAAGAAAGATTTGCTCAACAGTCAATCTTAGTT 1332
Db 322 AGACTATATGACTTTTGAAGTGTGGAACCAATCTGAGAAACAACCTATCTTGTG 381
Qy 1333 TTGATGAAGCTGATGCAATGTTGATATGAGTTTGTCTCAGAAATGAAGAAATT 1388
Db 382 CTGATGAAGCAGATGAATGCTTGAATGAGCTTTGAACCCCAATTAAGAAAGAT 437

RESULT 14
US-08-959-749-1
; Sequence 1, Application US/08959749
; Patent No. 5989858
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: McDevitt, Damien
; APPLICANT: Wang, Min
; APPLICANT: Traini, Christopher M.
; TITLE OF INVENTION: No. 5989858e1 Dbpb
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads

```

STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/959,749  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, Q. Todd  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: GM10113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215/994-2252  
TELEFAX: 215/994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-959-749-1

Query Match 3.9%; Score 84.6; DB 2; Length 1347;  
Best Local Similarity 53.5%; Pred. No. 1.3e-14;  
Matches 177; Conservative 0; Mismatches 154; Indels 0; Gaps 0;  
QY 1629 TATGCTTTGTGTAAGAACTAAGAAAGAGATTTTACTGCACTTTCTTGTCAAGA 1688  
DB 738 TATATTTTCTGTAATAGTAGAGATTAATGCAATGATTGACAGCTTCACTAATGAAGC 797  
QY 1689 AAAAATATCAACTACAACTATCCATGCTGATCGGAAACAGAGAGCGGAGCAAGCTCT 1748  
DB 798 TGGATTAATGAAGTTGGTATGATTCATGCTGTTAACCCACGTGAACGTAACAACTAAT 857  
QY 1749 TGGAGATTTTGGCTTTGGAAAGTGGCCAGTTCTTGTGTAATCTTCACTAGCTGCAGAGG 1808  
DB 858 GAAACGTAATCGTAATTTAGAAATTCATACGTTATGCGACGATTTAGCATCTCGTGG 917  
QY 1809 GCTGATATTTGAAATATGCAACATGTTATCAATTTTGAATCTTCTTCACTGATGTA 1868  
DB 918 TATTGATATTTGAAGGTGTTAGTATGATGATCAATTTTGAATGCTGCAATGATATGACTT 977  
QY 1869 ATATGTTCAATGCAATTTGGGCGTACTGCTGTTGGGAATATCGGACAGCAATTTCTT 1928  
DB 978 CTTTACCAATAGAGTTGGAGCAACTGCTGCGGGAATTAAGAGTGTAGCAATTAAGCT 1037  
QY 1929 TTTTGAATCTGGAATCGGATTAACCATTTAGCA 1959  
DB 1038 TTATAGTCTGATGAAGAACACAAATATTCA 1068

RESULT 15  
US-09-351-497-1  
Sequence 1, Application US/09351497  
Patent No. 6313270  
GENERAL INFORMATION:  
APPLICANT: Huang, Jianzhong  
APPLICANT: McDevitt, Damien  
APPLICANT: Wang, Min  
APPLICANT: Traini, Christopher M.  
TITLE OF INVENTION: No. 6313270e1 Dbpb  
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/351,497  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/959,749  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, Q. Todd  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: GM10113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215/994-2252  
TELEFAX: 215/994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-351-497-1

Query Match 3.9%; Score 84.6; DB 4; Length 1347;  
Best Local Similarity 53.5%; Pred. No. 1.3e-14;  
Matches 177; Conservative 0; Mismatches 154; Indels 0; Gaps 0;  
QY 1629 TATGCTTTGTGTAAGAACTAAGAAAGAGATTTTACTGCACTTTCTTGTCAAGA 1688  
DB 738 TATATTTTCTGTAATAGTAGAGATTAATGCAATGATTGACAGCTTCACTAATGAAGC 797  
QY 1689 AAAAATATCAACTACAACTATCCATGCTGATCGGAAACAGAGAGCGGAGCAAGCTCT 1748  
DB 798 TGGATTAATGAAGTTGGTATGATTCATGCTGTTAACCCACGTGAACGTAACAACTAAT 857  
QY 1749 TGGAGATTTTGGCTTTGGAAAGTGGCCAGTTCTTGTGTAATCTTCACTAGCTGCAGAGG 1808  
DB 858 GAAACGTAATCGTAATTTAGAAATTCATACGTTATGCGACGATTTAGCATCTCGTGG 917  
QY 1809 GCTGATATTTGAAATATGCAACATGTTATCAATTTTGAATCTTCTTCACTGATGTA 1868  
DB 918 TATTGATATTTGAAGGTGTTAGTATGATGATCAATTTTGAATGCTGCAATGATATGACTT 977  
QY 1869 ATATGTTCAATGCAATTTGGGCGTACTGCTGTTGGGAATATCGGACAGCAATTTCTT 1928  
DB 978 CTTTACCAATAGAGTTGGAGCAACTGCTGCGGGAATTAAGAGTGTAGCAATTAAGCT 1037  
QY 1929 TTTTGAATCTGGAATCGGATTAACCATTTAGCA 1959  
DB 1038 TTATAGTCTGATGAAGAACACAAATATTCA 1068

Search completed: June 10, 2003, 10:42:46  
Job time: 142 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2003, 09:43:39 ; Search time 296 Seconds  
(without alignments)  
10269.578 Million cell updates/sec

Title: US-09-714-865-15

Perfect score: 2172  
Sequence: 1 atgggggagatgaagattggga.....tagatgatgatcatgggat 2172

Scoring table:  
IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCITUS\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389.8	17.9	420	US-09-833-381-1475	Sequence 1475, App
2	386.4	17.8	2451	US-10-044-090-480	Sequence 480, App
3	384.8	17.7	4423	US-10-002-600-67	Sequence 67, App
4	317	14.6	591	US-09-764-891-1170	Sequence 1170, App
5	275.8	12.7	1902	US-09-938-842A-521	Sequence 521, App
6	220	10.1	2194	US-09-764-891-7481	Sequence 7481, App
7	220	10.1	2197	US-09-764-891-7482	Sequence 7482, App
8	220	10.1	2197	US-09-764-891-7483	Sequence 7483, App
9	210.2	9.7	3841	US-10-198-846-12716	Sequence 12716, A
10	168.4	7.8	403	US-09-960-352-1131	Sequence 14381, A
11	166.6	7.7	2010	US-09-964-824A-268	Sequence 268, App
12	166.6	7.7	2010	US-09-860-107-3397	Sequence 3397, App
13	164.4	7.6	2247	US-09-938-842A-1091	Sequence 1091, App
14	159.6	7.3	3501	US-09-938-842A-13	Sequence 13, App
15	158.2	7.3	2365	US-09-923-831-42	Sequence 42, App
16	153.6	7.1	1614	US-09-938-842A-2560	Sequence 2560, App
17	145.2	6.7	418	US-09-918-995-35717	Sequence 35717, A
18	136.6	6.3	1052	US-09-925-301-597	Sequence 597, App
19	136	6.3	1860	US-09-938-842A-699	Sequence 699, App

C	20	125.6	5.8	1830121	9	US-10-329-960-1	Sequence 1, App
	21	123.6	5.7	1830121	9	US-10-329-960-1	Sequence 1, App
	22	120	5.5	2202	9	US-09-938-842A-2296	Sequence 2296, App
	23	118.4	5.5	1575	10	US-09-815-242-9393	Sequence 9393, App
	24	117	5.4	1833	9	US-09-938-842A-1839	Sequence 1839, App
	25	115.2	5.3	3347	9	US-09-736-457-318	Sequence 318, App
	26	115.2	5.3	3347	9	US-09-902-941-318	Sequence 318, App
	27	115.2	5.3	3347	9	US-09-849-626-318	Sequence 318, App
	28	115.2	5.3	3347	9	US-10-017-754-318	Sequence 318, App
	29	112.6	5.2	2936	9	US-10-108-605-54	Sequence 54, App
	30	109.6	5.0	516	10	US-09-920-300A-784	Sequence 784, App
	31	109.6	5.0	516	12	US-10-033-528-784	Sequence 784, App
	32	109.6	5.0	551	10	US-09-920-300A-1344	Sequence 1344, App
	33	109.6	5.0	551	12	US-10-033-528-1344	Sequence 1344, App
	34	109.6	5.0	573	9	US-09-736-457-1384	Sequence 1384, App
	35	109.6	5.0	573	9	US-09-902-941-1384	Sequence 1384, App
	36	109.6	5.0	573	9	US-09-796-692-245	Sequence 245, App
	37	109.6	5.0	573	9	US-09-796-692-375	Sequence 375, App
	38	109.6	5.0	573	9	US-09-796-692-397	Sequence 397, App
	39	109.6	5.0	573	9	US-09-796-692-492	Sequence 492, App
	40	109.6	5.0	573	9	US-09-796-692-503	Sequence 503, App
	41	109.6	5.0	573	9	US-09-796-692-5064	Sequence 5064, App
	42	109.6	5.0	573	9	US-09-796-692-5261	Sequence 5261, App
	43	109.6	5.0	573	9	US-09-849-626-1384	Sequence 1384, App
	44	109.6	5.0	573	9	US-10-017-754-1384	Sequence 1384, App
	45	109.6	5.0	573	9	US-10-040-862-245	Sequence 245, App

## ALIGNMENTS

RESULT 1	US-09-833-381-1475	Application US/09833381
Sequence 1475	US-09-833-381-1475	
Patent No. US20020132090A1		
GENERAL INFORMATION:		
APPLICANT: Robison, Keith E.		
FILE REFERENCE: 5800-119		
CURRENT APPLICATION NUMBER: US/09/833,381		
CURRENT FILING DATE: 2001-04-11		
PRIOR APPLICATION NUMBER: 09/516,448		
PRIOR FILING DATE: 2000-02-29		
NUMBER OF SEQ ID NOS: 2050		
SOFTWARE: FastSeq for Windows Version 3.0		
SEQ ID NO 1475		
LENGTH: 420		
TYPE: DNA		
ORGANISM: Homo sapiens		
US-09-833-381-1475		
Query Match	17.9%	Score 389.8; DB 10; Length 420;
Best Local Similarity	99.5%	Pred. No. 2.2e-100;
Matches 391; Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
Qy	805	GACAAATACACACTATTCTTGTGAGAGTGTGACATGATGACACCAACCAATTCGTG 864
Db	28	GTCAAATACACACTATTCTTGTGAGAGTGTGACATGATGACACCAACCAATTCGTG 87
Qy	865	ACTTTTGAAGAGTAACTCTGTGTGACACACTGAATTAACAACATTTGTAAGCTGTTAT 924
Db	88	ACTTTTGAAGAGTAACTCTGTGTGACACACTGAATTAACAACATTTGTAAGCTGTTAT 147
Qy	925	ACTAAGTTACTCTGTGCAAAAAATAGATTCTCTCATCTTGAGAGACGAGATTG 984
Db	148	ACTAAGTTACTCTGTGCAAAAAATAGATTCTCTCATCTTGAGAGACGAGATTG 207
Qy	985	ATGGCTGTGCTCTCAACAGAGGTCTGGAGACACTGCGCTTTTCTTCAACCAATTTGGCT 1044
Db	208	ATGGCTGTGCTCTCAACAGAGGTCTGGAGACACTGCGCTTTTCTTCAACCAATTTGGCT 267
Qy	1045	CATTATGATGATGATGATGATTAAGAGTTTAAAGAGTTGACAGAACCAAGAGTGT 1104



FEATURE:  
NAME/KEY: misc: feature  
OTHER INFORMATION: Template ID: 350667.1  
US-10-002-600-67

Query Match 17.7%; Score 384.8; DB 12; Length 4423;  
Best Local Similarity 58.6%; Pred. No. 2,8e-98;  
Matches 762; Conservative 0; Mismatches 502; Indels 36; Gaps 4;

QY 788 AGACAGCAATTAACCTCGACAAATACGACACTATCTGTGGAAGTGTGACATGATG 847  
DB 536 ACACGGGAGTAACTTTGAGAAATATGATATATACAGTAGAGGCAACCGGACATGACT 595  
QY 848 CACACACAGCAATTCGACTTTTGAAGAACTATCTGTGACACATGAAATACACA 907  
DB 596 GTCCTCACATATTTAGAAATTTTAGCGATATGACATGGAGAAATATCATGGGGAACA 655  
QY 908 TTGCTAAAGTGTGATATCTAAGCTTACTCTGTGCAAAATATACGATTTCTATCATAC 967  
DB 656 TTGAACTTACTCGCTATCTGCTACTCTCACTGCAAGTGCAGAAACATGCCATTCCTATTA 715  
QY 968 TTGACAGACAGATTTGATGAGTGTGTGCTCAAA CAGGAGTGGGAGAGACTGGGCTTTTC 1027  
DB 716 AGGAGAAAGAGACTTATATGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775  
QY 1028 TCCTACCAATTTTGGCTCA-----TATGATGATGATG 1060  
DB 776 TTTTACCAATGAGTACAGATATATATACAGATGCTGCAAGAGAACTTTGAGGCTGTGA 835  
QY 1061 GAATTAACCTGCACTGCTTTTAAAGTGTGACAGAAACAGATGATTTATTTAGCACCA 1120  
DB 836 AGGAAATGAGAGATGAGGCGCGCGCAACATATCCAAATATCTTGTTTGGCCCCCA 895  
QY 1121 CTGAGAGATTTGTCACCAATTTATTTGAGCGAGAAATTTCTTTTGGAGCTGTG 1180  
DB 896 CAAAGAGATTTGCTGACATCTATGAGAGAGCAGAAATTTCTTACCGATCTAGAG 955  
QY 1181 TTAAGCTGTGTTTATATATGAGGAGAAACCACTGAGACATTCATTCAGCAATATGATAC 1240  
DB 956 TTGCTCTGTGATGATTTATGAGTGTGCTGATATGCTGACAGATTCGGGACTTGAAC 1015  
QY 1241 AAGGCTGTAATATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1300  
DB 1016 GTGAGATGCCACTTGTATGATGACCACTCCAGAGCTGATGATGATGATGATGATGATG 1075  
QY 1301 AGATTTGCTCAACAGATCAATATCTTATTTGATGAGAGCTGATGATGATGATGATG 1360  
DB 1076 AGATTTGATTAAGCTTCTGCAAGTACTTATGTTGATGATGATGATGATGATGATG 1135  
QY 1361 TGGGTTTGGTCCAGAAATGAGAAATTTCTTCCAGAGAAATTCATCAAGAGAAC 1420  
DB 1136 TGGGATTTGAACTCAGATACGTCGATATGTTGAACAAGATCTATGCCACCAAGGCG 1195  
QY 1421 AGGCGCAACCTTATGTCAGTGCAACTTTTCCAGAGAAATTCAAAGTTGGCTGAC 1480  
DB 1196 TTGCTCACACATGATGATTTAGTCTACTTTTCTTAAAGAAATTCAGATCTGCTGTG 1255  
QY 1481 AGTTTAAAGTCAATATCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1540  
DB 1256 ACTTTTGTGATGA--TATATCTTTTGTGCTGATGAGCAAGTATGCTTACCTGTGAGA 1312  
QY 1541 ATGTTACAGACCGCTTCTCAAGTTGGCAGTCTCAAAAAGAGAAAGCTGTTGAAA 1600  
DB 1313 ACATCACACAGAAAGTATGTTGGTGAAGACTTAAAGTCAATTTCTTACTGAGACA 1372  
QY 1601 TTCTGCAAAACATGAGG--GATGAAAGAACTATGCTTTTGTGAAATCAAGAAAAAG 1657  
DB 1373 TTTTAGTGAACAGAGGATGATCTTACTTATGATTTGTGAGAACCAAAAAAGGAG 1432  
QY 1658 CAGATTTTACGCACTTTCTTGTGCAAGAAAAATATCACTACAGATCCATGCTGTG 1717  
DB 1433 CAGATTCCTGAGAGATTTCTTATACATGAAAGATATGCTTGTACTAGTATCATGAG 1492

QY 1718 ATCGGAGACAGAGAGCGGAGACAACTCTTGGAGATTTTGGCTTTGGAAAGTGGCCAG 1777  
DB 1493 ACCGTCACAGAGAGATTCAGAGAGAGCCCTTACAGATTTCTGCTCAGGAAAAAGCCCA 1552  
QY 1778 TTCTTGTGCTACTTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1837  
DB 1553 TTCTAGGCTAGAGCTGTGGAGAGCAGAGAGCTTACATTTTAAATGTGAGCAATGTTA 1612  
QY 1838 TCATTTTGTATCTTCTTACATTTATGATGATGATGATGATGATGATGATGATGATGATG 1897  
DB 1613 TCATTTTGTATTTTCCAGATGATTTGAGAAATATGATGATGATGATGATGATGATGATG 1672  
QY 1898 GTTGTGGAATATGCGACAGCAATTTCTTTTGTATCTGATGATGATGATGATGATGATG 1957  
DB 1673 GTGTAGGAAACCTGGGCTTGGCCACCTCATTTCTTAA--TGAAAAAATGAAATATTA 1729  
QY 1958 CACAGGCTGATGAAGATTTAGACAGATGCTCAACAGAGATGTTCCGTGATGTTGGAAG 2017  
DB 1730 CAAAGATTTGTGATCTTTCTGTGAAAGCTTAAACAGAGAGTCTTCTTGTGGAAG 1789  
QY 2018 AAATGCTTTAGTACATACATTTCTGCTTCACTGATG 2057  
DB 1790 ATATGCTTATGACACACACTACAGAGTGGAGTGTGTG 1829

RESULT 4

US-09-764-891-1170  
; Sequence 1170, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1170  
; LENGTH: 591  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-1170

Query Match 14.6%; Score 317; DB 9; Length 591;  
Best Local Similarity 99.4%; Pred. No. 1.4e-79;  
Matches 317; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1854 TTCTACCATTTGATGATATGTTGATGATGATGATGATGATGATGATGATGATGATG 1913  
DB 1 TTCTACCATTTGATGATATGTTGATGATGATGATGATGATGATGATGATGATGATG 60  
QY 1914 CAGAGCAATTTCTTTTGTGATCTTGAATGGAATACATTTAGCACAGCTCTAGTAA 1973  
DB 61 CAGAGCAATTTCTTTTGTGATCTTGAATGGAATACATTTAGCACAGCTCTAGTAA 120  
QY 1974 AGTATGACAGATGCTCAACAGATGCTTCTGATGATGATGATGATGATGATGATGATG 2033  
DB 121 AGTATGACAGATGCTCAACAGATGCTTCTGATGATGATGATGATGATGATGATGATG 180  
QY 2034 ATACATTTCTGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2093  
DB 181 ATACATTTCTGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
QY 2094 AAAAGGACAGACATTTTGAACACAGCTGCTTTTCTTCAAGCTTCCATCCAGT 2153  
DB 241 AAAAGGACAGACATTTTGAACACAGCTGCTTTTCTTCAAGCTTCCATCCAGT 300  
QY 2154 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2172  
DB 301 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 319

RESULT 5  
 US-09-938-842A-521  
 ; Sequence 521, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Krepes, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE  
 ; FILE REFERENCE: SCRIPI300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938,842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 521  
 ; LENGTH: 1902  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-521

Query Match 12.7% Score 275.8; DB 9; Length 1902;  
 Best Local Similarity 52.8%; Pred. No. 1.6e-67;  
 Matches 707; Conservative 0; Mismatches 602; Indels 30; Gaps 4;

QY 698 GAGAGAAAGTAGATGATCTCAAGACCAAGAGTACATACATACCCCTCCCTCACTG 757  
 DB 311 GAAGCGAGGAGTGGACCGTAGGAGACGTGAAGTAAACCCCTTGAAAGATGATTCG 370  
 QY 758 AGATGAGAGCTCCATCTTTGACATTATGACAGGACATTAACCTTGACAAATACGACA 817  
 DB 371 AACAGAACAGCTTTTACTAGAGAGATTAATACGTTAATTTTGAATGCTATAG 430  
 QY 818 CTATCTTGTGAGAGTCTGAGCATATGACACCAAGCATTTCTACTTTGAAAG 877  
 DB 431 ATATTCGATTTGAACACGAGTGGGATATGCTCTCTCTTTAACACATTTGACAGAGA 490  
 QY 878 CTATCTCTGTGACAGACTGATTAACAACTGCTAAAGCTGTTACTAAGCTTCTC 937  
 DB 491 TAGATCTCGGAGGAGCTTGAATTAATCCGTATGCAATTAATGTTAAAGCAAC 550  
 QY 938 CTGTGCAAAATATCAGATTTCTATCATCTTGCAGAGACGAGATTTGATGCTGCTC 997  
 DB 551 CTGTTCAGCGTCAATGCGATTCGATTTGCTTGAAGGAGGATTTGATGCGCTGCTC 610  
 QY 998 AACAGAGGCTCGGAGAACGCGGCTTTCTCCCAACAAATTTGGCTCATATGATGATG 1057  
 DB 611 AGACAGGCTCGGAGAACAGCTGCTTTTGTCTTCAATCATTAAGGAAATATGAAG 670  
 QY 1058 ATGGAATTAATCGCAG---TCGTTTAAAGATTGCAAGAACGAGATGATTTAG 1114  
 DB 671 ATCAGCATGTAACAGACCCCGTGTTCACGAACAGTCTACCTCTTGCACTTATCTCT 730  
 QY 1115 CACCACTCGAGATTTGTCACCAAGATTTATTTTGAAGCCAGAAATTTCTTTTGGGA 1174  
 DB 731 CACCAACAGAGAGTGGCAAGTCAATGATGATGAGGCTTAAAGATTTCTTTACCAA 790  
 QY 1175 CTGTGTGAAGAGCTTTGTTATATATGAGGGAACCCAGCTGGGACATTCATTTGACAAA 1234  
 DB 791 CTGTGTGAAGAGTGTGTGTCATATGAGGAAACCTATTAAACACAGCTCGGGAGC 850  
 QY 1235 TAGTACAGAGGCTGAATATATATATGCTACTCTCGAAGAGCTAGTATATCATAGCA 1294  
 DB 851 TTGAGAGGAGAGTGTATATTTCTGTGCAACGCTGTGCTTAAATGATTTGCTCGAGA 910  
 QY 1295 AAGAAAGATTTGCTCAACAGATCAATTAATAGTTTGGATGAAGCTGATGATGT 1354

DB 911 GAGCTAGAGTCTCAATGACATGATTAATTTTACCTCTTGATGAGCGCATAGATGC 970  
 QY 1355 TGGAATGAGTTTGGTCCAGAAATGAAGATTAATTTCTGGCCAGGAATGCCATCA 1414  
 DB 971 TTGACATGGGTTTGAACCAATTAATGAAGATTTGCAACAAATGACATCTCCAC 1030  
 QY 1415 AGGACAGGCGCAACCCCTTATGTTGCAAGTCACTTTCCAGAGAAATTCAAAGGTTGG 1474  
 DB 1031 GTGAGATTAGACAGACACTGTTGTTAGTCAATTTCCAGAGAAATTCAGAGACTCG 1090  
 QY 1475 CTGCAAGTTTAAAGTCAATTAATCTGTTGCTGTTGCAAGAGTGGAGCAT 1534  
 DB 1091 CAGCTACTTCCAG---CAAAATATATATTTTGGCTGTGGTAGAGTGGCTTCAAGTA 1147  
 QY 1535 GTAGAGATGTTGACAGACCGTCTCCAGTGTGGCCAGTTCTTCAAAAGAGAAAGCTCG 1594  
 DB 1148 CCGATTAAATTTGTCCAAAGGTTGAATTTGCTTACTCTGACAAAGAGCATCTCA 1207  
 QY 1595 TTGAAATTTGCGAAACATAGGGATGAAGA-----ACTATGG 1633  
 DB 1208 TGAGCTGCTTACCGCTCAGAGAGAGATGGATCCAGGCAAGCAAGCCCTGACCTTGA 1267  
 QY 1634 TCTTTGTTGAAGTAAAGAAAGAGCATTTTACTGCAACTTTCTTTGCAAGAAAAA 1693  
 DB 1268 TTTTGTGAGACAAAGAGAGAGTGAAGTCTTGTGAAATTTGTTGTCATCAATGGGT 1327  
 QY 1694 TATCAACTCAAGATTCATGATGATGATGCGGAGACAGAGAGCGGAGCAAGCTCTTGAG 1753  
 DB 1328 TTCCACCAACCTTCATTCACGTTGACAGAACACAGAGAAAGAGATGGCATTTGAAAG 1387  
 QY 1754 ATTTTGTGTTGAAAGTCCCGCAGTTCTTGTGCTACTTCAATAGTGCAGAGGCTGG 1813  
 DB 1388 CTTTCAAGAGTGGGAACACCGATTTGTTGTTGCACTGATGATGACAGCAGTGGCTTG 1447  
 QY 1814 ATTTGAAATGTCACACTGTTATCAATTTGATCTTCTTACATTTGATGAATATG 1873  
 DB 1448 ACAATTCACAGCGGTCAATGTTGTCACCTTGTGCAATGATGATGATGATGATG 1507  
 QY 1874 TTGATGAATTTGGGCGTACTGCTGTTGTGGGAATCTGCGACAGGCAATTTCTTTTGG 1933  
 DB 1508 TCCACCGTATTTGACCAACAGACGCTGCGCAATCAGAGACTGACATCGCTTTTCA 1567  
 QY 1934 ATCTGAATCGATTAACATTTAGCAACAGCTCTTGAATAAAGTATTAAGATGCTCAAC 1993  
 DB 1568 A---CGATGCAACACTCTACTGCGCAGACCGCTGCTGATGATGATGATGATGATG 1624  
 QY 1994 AGGATGTTCTCGATGCTT 2012  
 DB 1625 AAGAGTCCCTGATGCT 1643

RESULT 6  
 US-09-764-891-7481  
 ; Sequence 7481, Application US/09764891  
 ; Publication No. US20030077808A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC006  
 ; CURRENT APPLICATION NUMBER: US/09/764,891  
 ; CURRENT FILING DATE: 2001-01-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 10231  
 ; SOFTWARE: Patent Ver. 2.0  
 ; SEQ ID NO 7481  
 ; LENGTH: 2194  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-891-7481

Query Match 10.1% Score 220; DB 9; Length 2194;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-51;  
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Oy	1131	AGTAAACAGATTATTTTGAAGACCGAAAAATTTCTTTGGGACTGTGTGAAGAGCTGT	1190
Db	867	GGCCCAACAGGTGCAGCAAGTNGCTGCTGAATATTTGTAGACATGTCCCTGAAGCTTAC	926
Oy	1191	TGTTATATATGAGGAGAACCCAGCTGGGACATTCAAATTCGACAAATATGTAACAAGCTGTAA	1250
Db	927	TTGTATCTACGGTGTGCTCCTTAAGGGACCAACAATATCGTATTTGGAGAGAGGTGTGA	986
Oy	1251	TATATATNGTCTACTCCTGGAAAGCTATGATATCATATGGCAAAAGAAAGATTGGTCT	1310
Db	987	AATCTGATTCACAACCTGGAAAGCTATTTGACTTTTAAAGTGTGAAAAAACCAATCT	1046
Oy	1311	CAAAACAGATCAAAATCTAGTTTGGATGAAGCTATGSCATGTTGGATATGGGTTTGG	1370
Db	1047	GAGAAAGAACACCTTACTGTTCCTTGAATGAACAGATAGAAATCTGTGATATGGGCTTGA	1106
Oy	1371	TCCAAAATGAGAAGATTATTTCTTGGCCAGAAATGTCATCAAGAACACGCGCAAC	1430
Db	1107	ACCCCAAAATAGGAAGATTGTGAT-----CAATATAGACTGTATAGCAAC	1154
Oy	1431	CCTATATGTCAGTGAACTTTTCAGAGAAATTCAAAGTTGGTGGAGACTTTTAA	1490
Db	1155	TCTATATGAGTGTGACTGTGGCCAAAAGAAATGAACAAGCTGTGAAGATTTCCTAA	1214
Oy	1491	GTCAAAATATCTGTTTGTGTGTGCTGTGGAACAAGTGGTGGAGCATGTAGATGTTCACGA	1550
Db	1215	AGACTATATTCATATAAACATTGGTGACCTTGAACATGATGTGAACCAACCAATTTCTTCA	1274
Oy	1551	GACCTTCTCCAAAGTTGGCCAGCTTCCAAAAGAGAAAAGCTCGTTGAAATTTTCGAA	1610
Db	1275	GATTGTGATGTGTCTCATGACGTGAAGAAAGATATAAAACTATTTCTCTAATGGAAGA	1334
Oy	1611	CATAGGGATG-----AAAGAACTATGTGCTTTGTGTGAATTAAGAAAAAGCAGA	1661
Db	1335	GATCATGATGAGAAAGAAATAAACCATTTGTTTTGTGAAACCAAAAGAAATGTGA	1394
Oy	1662	TTTTACTGCACCTTTCTTGTTCAGAAAAAATATCAACTACAGTATCCATGTGATCG	1721
Db	1395	TGAGCTTACGAAAAATAGAGAGAGATGGGTGGCTCGACATGGGTATCCATGTGTGAA	1454
Oy	1722	GGAACAGAGAGAGCGGAGCAAGCTCTTGGAGATTTTGGCTTTGAAAGTCCCAAGTTCT	1781
Db	1455	GAGTCAACAAAGCGCTGACTGGGTTCTAAATGAATTTCAACATGGAAAAAGCTCCATATCT	1514
Oy	1782	TGTTGCTACTTCAGTAGCTGCAGAGGGCTGGATTTGAAAAATGCAACAATGTATCAA	1841
Db	1515	GATTGCTACAGATGTGGCTCTCCAGAGGGCTAGATGTGAAATGTGAATTTGTTCATCAA	1574
Oy	1842	TTTTGATCTTCTTCTACCATTTGATGAATATGTTTCATCGAATTTGGGCTACTGTCGTG	1901
Db	1575	TTATGACTAACCTTAACTCTCAGAGAGATTAATTCATCGAATTTGGAAAGACTGCGCAG	1634
Oy	1902	TGGGAATACTGGCAGAGCAATTTCTTTT 1932	
Db	1635	TACCAAAACAGCAGACATACCTTTCTTT 1665	

```

RESULT 10
US-09-960-352-14381
; Sequence 14381, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nenping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nageshan
; TITLE OF INVENTION: NOCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
;

```

```
; SEQ ID NO 14381
;
; LENGTH: 403
;
; TYPE: DNA
;
; ORGANISM: Bos taurus
;
; OTHER INFORMATION: Clone ID: 61-LIB34-050-Q1-E1-H2
US-03-960-352-14381
```

Query Match	7.8%;	Score 168.4;	DB 10;	Length 403;
Best Local Similarity	64.8%;	Pred. No. 2e-37;		
Matches 250;	Conservative 0;	Mismatches 136;	Indels 0;	Gaps 0;

QY	1112	TAGACACCACTCGAAGAAATGGTCTCAACCGATTATTTGGAGCCAGAAAATTTCTTTTG	1177
Db	14	TAGACACCACTAGAGAAATTTGGCCGTACAGATCTATAGAGAAAGCCAGAAAATTTTTCATACCC	73
QY	1172	GGACTTGTGTAGAGACTGTGTTGTATATATAGGGGGAAACCCAGCTGGGACATTTCAATTGGAC	1231
Db	74	GATCTAGAGTTCTGCTCTTGTGTGGTTATATGGTGGTGCTGATATTGGTCAAGCAAAATTCGAG	133
QY	1232	AAATATGTCAAAGCTGTATATATTTATGTGCTACTCTCGGAAAGACTATGATATCATNG	1297
Db	134	ACTTAGAACCGTAGATGCCATTATTAATAGTGGACCTCCAGGACGCTTAATGGAATATGATGG	193
QY	1292	GCAAGAGAAAAGATTGGTCTCAAAACAGATCAATATCTTAGTTTGGATTAACGTGATCGCA	1351
Db	194	AAAGAGGAAAAATTTGATTAAGATCTTCGCAAGTACTTGGTGTATAGATTAAGCCGATCGGA	253
QY	1352	TGTTGGATATGGGTTTGGTCCAGAAAATGAAGAATTAAATTTCTTGCCACAGAAATGCCAT	1411
Db	254	TGTTGGATATGGGTTTGAACTCAGATACGTAGAAATGTTGGAACAAAGATCTACTATGCCAC	313
QY	1412	CAAGGAAACAGCGCCAAACCTTATATGTTCAGTGCAATCTTTCAGAGGAATTTCAAAGT	1471
Db	314	CAAGGGAGTTTGGCACACTATGATGTTTAACTGTACTTTCCTTAAGGAAATACAGATGC	373
QY	1472	TGGCTGCAGAGTTTTTAAAGTCAAT	1497
Db	374	TTGCTCGTGAATCTTGGATCAATAT	399

```

RESULT 11
US-09-964-824A-268
; Sequence 268, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signature
; TITLE OF INVENTION: Secs
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236, 033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 268
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-268

```

	Query Match	Similarity	Score	DB	Length
Best	Local	50.0%	166.6	10	2010
Matches	551	Conservative	0	Mismatches	519
				Indels	33
				Gaps	4

  

	Query	Score	DB	Length
Db	838 GGACATGATGCACCCAGCAATTCGTGACCTTTGAAGAAGCTAATCTCTGCACACTG	897		
	250 GGAATGTTTTCCTAAACCGGTGTTCCTTCATCATGCTTAACCTTCCCAATATATGA	309		

QY 898 AATAACAATCTTCTAAAGCTGGTATATCTAAGCTTACTCTGTCGAAAAATACAGATT 957  
Db 310 ATGAGATGTGTGATGATCAGCACTTTACAGAAACCAATCCAAATTCAGTCCAGGATTT 369  
QY 958 CCTATCATCTTACAGAGAGATTTGATGGCTTTGGCTCAAAACAGGCTCGGAAGACT 1017  
Db 370 CCGTTGGCTTTTACGCGGGAATGATGGGATGCGATGCTGACAGCTGCTCTGGGAAGACG 429  
QY 1018 GCGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGGAATTAATGCGACGCT 1077  
Db 430 TTGGCGATCTCTGCTGCAATTTGTCATATTAACCAACAGCCATC----- 477  
QY 1078 TTTAAAGATTGAGAGAACCAAGATGATATTTGTAGCAACCACTGAGATTGTCAC 1137  
Db 478 TTGAAAGGGGAGATGGCCCAATCTGCTAGTTCTGGCTCTACAGAGAGCTTGCCAG 537  
QY 1138 CAGATTTATTTGAGAGCAAAAATTTCTTTGGGACTGTGTAGAGCTGTGTTATA 1197  
Db 538 CAGATACAGAGATGGCCGATGACTATGAGCAAAATGTTCTAGATTGAAGACTCTGTATT 597  
QY 1198 TATGGGGAAACCCAGCTGGAGCACTTCAATTCAGCAAAATAGTACAGGCTGTAATATTA 1257  
Db 598 TATGAGAGTGTCTCTAAAGTCTCCAGATTCGAGACTTGGAAAGAGGTGTGAGATGTC 657  
QY 1258 TGTGCTACTCTGAGAGACTGATGATATCATAGGCAAGAAAGATGCTTCAACAG 1317  
Db 658 ATAGCCACTCTGAGACGCTGATAGATTCTCTGAGTCAAGAAAGCAAAATCTTCGCGA 717  
QY 1318 ATCAAAATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377  
Db 718 TGTACTTACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777  
QY 1378 ATGAGAGATTAATTTCTCCAGAGATGCTCAAGAGAGAGAGAGAGAGAGAGAGAG 1437  
Db 778 AT-----CGTAAATTTGTTACCAATCAGAGCTGATAG-----CAGACACTGATG 825  
QY 1438 TTGAGTGAACCTTTTCCAGAGAAATTCAAAGTTGGCTGACAGATTTTAAAGTCAAT 1497  
Db 826 TGGAGTGAACCTTTGCGAAAGATGATGAGAGCTTCAAGAGATTTCTTCGATATAC 885  
QY 1498 TATCTGTTTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557  
Db 886 ACCCAGATCAAGATGAGCAATCTGAGATGATGATGATGATGATGATGATGATGATG 945  
QY 1558 CTCGAATGGCAGATCTCAAAAAGAGAAAGCTGTTGAAATTCGCAAAATGAGG 1617  
Db 946 GATGCTGATGAGAAAGTGAAGAACCAAGATTCATCAATGATGAGAAATATG 1005  
QY 1618 GATGAA-----AGAACTATGCTTTGTTGAACTAAGAAAGAGAGATTTACT 1668  
Db 1006 GTGAAAGAGAAACAAATTAATTTGTGAGACAAAGAGAGCTGTGATGATCTG 1065  
QY 1669 GCAACTTTTCTTTGTCAAGAAAAATTCACATCAAGTATCCATGATGATCGGAAACAG 1728  
Db 1066 ACTGGAAGATGCGCAGAGATGTTGGCAGCTATGATCAATGATGAGCAAGATCAA 1125  
QY 1729 AGAGAGCGGAGAGAGCTTTGAGATTTTGGCTTTGAGAGAGAGAGAGAGAGAGAGAG 1789  
Db 1126 CCGAAAGAGATTTGGTACTTAATGATTCGTTCTGGAAGAGAGAGAGAGAGAGAGAG 1185  
QY 1789 ACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848  
Db 1186 ACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1245  
QY 1849 CTTCCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1908  
Db 1246 TATTCAAAGAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305  
QY 1909 ACTGAGAGAGAGATTTCTTTT 1931  
Db 1306 AAGGATACCGCTTACTTCTT 1328

RESULT 12  
US-09-880-107-3397  
; Sequence 3397, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Schoerf, Joseph G.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3397  
; LENGTH: 2010  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 US9321  
US-09-880-107-3397

Query Match 7.7%; Score 166.6; DB 10; Length 2010;  
Best Local Similarity 50.0%; Pred. No. 1,9e-36;  
Matches 551; Conservative 0; Mismatches 519; Indels 33; Gaps 4;

QY 838 GCAGATGATGACACCAAGCAATTCGACTTTTGAAGAGTAACTCTGTCAGACATG 897  
Db 250 GAGATGTTGTCTTAAACCGGTTTGTCTTCCATCATGCTACTTCCACATATGTA 309  
QY 898 AATAACAATCTTCTAAAGCTGGTATATCTAAGCTTACTCTGTCGAAAAATACAGATT 957  
Db 310 ATGAGATGTGTGATGATCAGCACTTTACAGAAACCAATCCAAATTCAGTCCAGGATTT 369  
QY 958 CCTATCATCTTACAGAGAGATTTGATGGCTTTGGCTCAAAACAGGCTCGGAAGACT 1017  
Db 370 CCGTTGGCTTTTACGCGGGAATGATGGGATGCGATGCTGACAGCTGCTCTGGGAAGACG 429  
QY 1018 GCGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGGAATTAATGCGACGCT 1077  
Db 430 TTGGCGATCTCTGCTGCAATTTGTCATATTAACCAACAGCCATC----- 477  
QY 1078 TTTAAAGATTGAGAGAACCAAGATGATATTTGTAGCAACCACTGAGATTGTCAC 1137  
Db 478 TTGAAAGGGGAGATGGCCCAATCTGTAGTTCTGCTCTACAGAGAGCTTGCCAG 537  
QY 1138 CAGATTTATTTGAGAGCAAAAATTTCTTTGGGACTGTGTAGAGCTGTGTTATA 1197  
Db 538 CAGATACAGAGATGGCCGATGACTATGAGCAAAATGTTCTAGATTGAAGACTCTGTATT 597  
QY 1198 TATGGGGAAACCCAGCTGGAGCACTTCAATTCGCAAAATAGTACAGGCTGTAATATTA 1257  
Db 598 TATGAGAGTGTCTCTAAAGTCTCCAGATTCGAGACTTGGAAAGAGGTGTGAGATGTC 657  
QY 1258 TGTGCTACTCTGAGAGACTGATGATATCATAGGCAAGAAAGATGCTTCAACAG 1317  
Db 658 ATAGCCACTCTGAGAGCTGATAGATTTCTGAGATGAGAAAGAGCAAAATCTTCGCCGA 717  
QY 1318 ATCAAAATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1377  
Db 718 TGTACTTACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777  
QY 1378 ATGAGAGATTAATTTCTTCGAGAGATGCAATCAAGAGAGAGAGAGAGAGAGAGAG 1437  
Db 778 AT-----CGTAAATTTGTTACCAATCAGAGCTGATAG-----CAGACACTGATG 825  
QY 1438 TTGAGTGAACCTTTTCCAGAGAAATTCAAAGTTGGCTGACAGATTTTAAAGTCAAT 1497

Db 826 TGGAGTCAACTGGCCAAAAGTAAGACAGCTTCAGAGATTTCTTCGTGATAC 885  
Qy 1498 TATCTGTTTGTCTGTGGAACAAGTGGTGAGCACTGATGATGTTCAAGCAACCTT 1557  
Db 886 ACCCAGATCAACGATGAGCAATCTGAGAGTGAAGCCAAACATCTCCAGATAGTG 945  
Qy 1558 CTCGAAGTTGGCCAGTTCTCAAAAAGAGAAAGCTCGTTGAAATTCGCAAAACATAGG 1617  
Db 946 GATGTCGATGAGAAAGTGAAGAAAGACCAAGATGATCCAACTAATGAGAAATATAG 1005  
Qy 1618 GATGAA-----AGAACTATGCTTTTGTGAAAATAAGAAAAAGCAATTTTACT 1668  
Db 1006 GCTGAAAAGGAAAAAACAATTAATTTGTGAGACAAAGAGACCTGTGATGATCTG 1065  
Qy 1669 GCAACTTTTCTTGTCAAGAAAAAATATCAACTACATGATCCATGATGATGAGAAAG 1728  
Db 1066 ACTGGAAGATGGCAGAGATGTTGGCCAGCTATGTTATCCATGAGACAAAGAGTCAA 1125  
Qy 1729 AGAGAGCGGAGCAAGCTCTTGAGATTTTGTGTTGAAAGTCCAGTTCTTGTGCT 1788  
Db 1126 CCAGAAAGATTTGGTACTTAATGAGTTCGTTCTGAAAAGGACCCATCTTATTTGCT 1185  
Qy 1789 ACTTCACTAGCTGCGAGAGGCTGATATTTGAAAATGTCACATGTTATCAATTTTAT 1848  
Db 1186 ACAGATGATGCTCAGCTGGGCTGATGTGAAGATGTCAAGTTTGTGATCACTATGAC 1245  
Qy 1849 CTTCCTTCTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1908  
Db 1246 TATCCAAACAGCTCAGAGATTAATGTGACCGTATTTGGCCGAACGCGTGAAGCAAC 1305  
Qy 1909 ACTGCGAGAGCAATTTCTTTT 1931  
Db 1306 AAGGTACCGCTATACCTTT 1328

RESULT 13  
US-09-938-842A-1091  
; Sequence 1091, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kieps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; CURRENT APPLICATION NUMBER: US/09/938, 842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227, 866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264, 647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300, 111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1091  
; LENGTH: 2247  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1091

Query Match 7.64; Score 164.4; DB 9; Length 2247;  
Best Local Similarity 49.88; Pred. No. 8.5e-36;  
Matches 600; Conservative 0; Mismatches 571; Indels 33; Gaps 6;

Qy 832 GTGCTGACATGATGACCAACCAATCTGACTTTTGAAGAGCAATCTCTGTCAG 891  
Db 652 GTCTCGGTTTGTATGTTCAATGAGCACTTAAGACCTTTGAGATTTGATTTCTCA 711  
Qy 892 AACTGATTAACAACATTTGCTAAAGCTGTTATATCTAGCTTACTCTGTGCAAAATAC 951  
Db 712 CAGATTAAGAGTGTATCAAAAGCAAGCTTATGAAAAAAGCTTACAGCAATCTCAATGTAC 771

Qy 952 AGATTCCTATCATACTTGCAGAGACGATTTGATGCTTGTCTCAAAACAGGCTGCG 1011  
Db 772 GCTTACCTATGTTGTGCTGTGTCAGATGTTATTTGGCATAGCCAAACCTGTTACAGT 831  
Qy 1012 AAGACTGCGGCTTTTCTCTACCAATTTTGGCTCATATGATGATGATGAATTAAGTCC 1071  
Db 832 AAGACTGCGGCTTTTCTCTCAATGATTTGATTAATGATGATGATGATGATGATGAT 891  
Qy 1072 AGTCTTTTAAAGTTCAGAAACCAAGTGTATATGATGATGATGATGATGATGATGAT 1131  
Db 892 AGAGAT-----GAAAGTCAATGTTGTTATGTTGTTGTTGTTGTTGTTGTTGTT 939  
Qy 1132 GTCACCAATTTATTTGAGAGCAAAATTTTCTTTTGGCACTTGTGTAAGCTGTT 1191  
Db 940 GCTCATCAATCTCTTGAAGCCAAAGATTTTCAAAAGCATTTGCTTACAGAGCTCTG 999  
Qy 1192 GTTATATATGAGGGAACCAAGCTGGACATTCATTCGACAAATATGATACAGGCTTAAT 1251  
Db 1000 GCTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1059  
Qy 1252 ATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311  
Db 1060 ATATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119  
Qy 1312 AAACAGATCAAAATCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1371  
Db 1120 ATGAGAGCCAGTTATTTGTTGTTGATGAGGAGATGATGATGATGATGATGATGATGAT 1179  
Qy 1372 CCAGAAATGAGAGTTAATTTCTTTCGAGAGATGATGATGATGATGATGATGATGATGAT 1431  
Db 1180 CCAGAGT-----AAGGTATATGTTGCGAGATTTCTCT-----GATCGCCAGACT 1227  
Qy 1432 CTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491  
Db 1228 TTAATATTTTCAAGCCATATGCTTGAAGATGATGATGATGATGATGATGATGATGAT 1284  
Qy 1492 TCAATTTATCTGTTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551  
Db 1285 TCAATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344  
Qy 1552 ACCGT---TCTCCAGTGGCCAGTTCTCAAAAAGAGAAAGCTCGTTGAATTTGCGA 1608  
Db 1345 GTTGTCAATGCTATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404  
Qy 1609 AATATGAGGATGAAGAACTATGCTTGTGATGATGATGATGATGATGATGATGATGAT 1665  
Db 1405 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464  
Qy 1666 ACTGCAACTTTTCTTGTCAAGAAAAATATCACTACATGATGATGATGATGATGATGATGAT 1725  
Db 1465 ATGGAAGCTCACTTACTCTTAATTTCTTCAAGTTGCTGCTTTACGCTGATTAAGAC 1524  
Qy 1726 CAGAGAGAGGAGGAGCAAGCTCTTGAGATTTTGTGTTGAGAAAGTCCAGTTCTTGT 1785  
Db 1525 CAAGCATCAAGATGAGAACTCTGAGAAAGTTCAAAATCTGAGTTCAACATGTCGAT 1584  
Qy 1786 GCTACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845  
Db 1585 GGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1644  
Qy 1846 GATCTTCTTCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1905  
Db 1645 GATCTGCAAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1704  
Qy 1906 AATATGAGCAAAATTTCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1965  
Db 1705 GACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764  
Qy 1966 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2025  
Db 1765 TTGTTAAACAGTGTGTTGCTGCTGATCAAAATGATGATGATGATGATGATGATGATGAT 1824



QY 2026 TTTA 2029  
 Db 1825 ATGA 1828

RESULT 14  
 US-09-938-842A-13  
 ; Sequence 13, Application US/09938842A  
 ; Patent No. US20020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE  
 ; FILE REFERENCE: SCRIPI300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 13  
 ; LENGTH: 3501  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-13

Query Match 7.3%; Score 159.6; DB 9; Length 3501;  
 Best Local Similarity 48.5%; Pred. No. 2,7e-34;  
 Matches 589; Conservative 0; Mismatches 589; Indels 36; Gaps 4;

QY 827 TGAAGTGTGACATGATGACACCAACCAATTCGACTTTGAAGAAGTAACTCT 886  
 Db 1550 TGAAGTTCATGAGAAAGATGTTCCAAAGCCCAATTAAGTTGGCACCAGACTGGCTAA 1609

QY 887 GTGAGACATGAAATCAATATGCTTAAGCTGTATATAGCTTACTCCCTGTGAAA 946  
 Db 1610 CAGCAAAATTTGGAGTACTATGAGAAAGCTTAATATGAAAAGCCAAATCCCTAATCAA 1669

QY 947 AATACAGTATTCCTATCATCTTGACAGAGCAATTTGATGCTGTGCTCAAGAGGT 1006  
 Db 1670 CACAGAGACTCCCTATCATCATGAGCGGTGACAGCTGATGGGGTTGCAAAAACGGTT 1729

QY 1007 CTGGGAAGACTGGGGCTTTTCTCTACCAATTTGGCTCATATGATGATGAAATTA 1066  
 Db 1730 CAGGTAAACCTTGGTTGTTTGGCAATGCTGAGCAATTTAAAGATCAGCC----- 1784

QY 1067 CTGCCAGCTCTTTTAAAGATGACAGAACCAAGTATTTATTTGACCACTCGAG 1126  
 Db 1785 -----TCCGTTGAAGCTGTGATGGCCAAATGCTCTTGTAAATGACCTACTAGGG 1837

QY 1127 AATTGTCACACAGATTTATTTGGAAGCCAGAAAATTTCTTTTGGGACTGTGTAAG 1186  
 Db 1838 AGCTGTTCAGCAGATTCACAGGATATCGAAAAGTTTCTTAAGCCATTGGGTAATAGAT 1897

QY 1187 CTGTGTTATATATGAGGGAACCCAGCTGGGACATTCGAATTCACAAATAGTCAAGGCT 1246  
 Db 1898 GCGTCCCTGTGATGAGGATCAGAGTTGCCCGCAAAATTAAGTAAGTAAAGGTA 1957

QY 1247 GTAAATATATATGCTACTCTCTGGAAGACTGATGATATCATATAGCAAG----- 1297  
 Db 1958 CTAGATTTGTGTGTCACCTCTGGGGAATGATGATATCTTTTGGACAGAGTGGGA 2017

QY 1298 AAAAGATTTGCTCAACAGATCAATATAGTTTGGATGGAAGCTGAGATGTTGG 1357  
 Db 2018 AAATCACTATCTGCGAAGATCAATTTTGGTAATGATGAAAGCTGATGATGTTG 2077

QY 1358 ATATGGGTTTGTTCAGAAATGAAAGTTAAATTTCTTCCAGAGATGCATCAAGG 1417

Db 2078 ACATGGGTTTGAAGCCTCAAAATTAATCTGTATTAATTCAGAACTTCACCG----- 2127

QY 1418 AACAGGCCCAACCCCTTATGTCAGTCAACTTTTCAGAGAAATTCAAAGTTGGCTG 1477

Db 2128 --GAAGCCAAACTGTGCTTTTCTTCTGCACTTTTCAGCCCAAGTTGAATTTGGCAC 2185

QY 1478 CAGAGTTTTTAAAGTCMAATTAATCTGTTTGTCTGCTGTTGACAGAGTGGGTGAGCATGA 1537

Db 2186 GTAAAGTCTTGAACAAAGCTGTAGAAATACAGGTGGTGGAGAGATGTTGTGAATAAAG 2245

QY 1538 GAGATCTTCAGCAGACCGCTTTCCAAAGTTGGCCAGTTCTCAAAAAGAAAGCTGTTG 1597

Db 2246 AATAAATCAGTATGATTAAGTCAGACCCGAGAGATGATAGTCTTCTTACACTTCGGAAC 2305

QY 1598 AATTTGCGAAACATAGGGATGAAAGAACTATGCTCTTGTGAACTTAAGAAAAG 1657

Db 2306 TTCTTGGGAAATGTTGTGAAAGGAAATTTCTGTTCTTTGTTCACTGTCAGGAAAAAT 2365

QY 1658 CAGATTTTACTGCACTTTTCTTGTCAAGAAAATATCAACTACAGTATTCATGCTG 1717

Db 2366 GTGATGCCCTGTATAGGATATGATTAATCTAGTACCTTGTCTGTCTTTCACAGGG 2425

QY 1718 ATCGGAACAGAGAGAGCGGAGCAAGCTCTTGAGATTTTGTGTTGAAAGTCCAG 1777

Db 2426 GTAAGATCAGACCGATCGTGAATCACTATATCTGATTTTAAAGCATGCTGCAATT 2485

QY 1778 TTCTTTGCTACTTCTAGTATGCTGCCAGAGGGCTGATATTAATAATGTCACAACTGA 1837

Db 2486 TGTGATTCACCAAGTGTGCTGCTGAGGGGTCTTAATATGAAAAGCTTGAGTTGGT 2545

QY 1838 TCATATTTGATCTCTCTTCTTACCATGATGATATATGTCATGCAATTTGGCGTACTGTC 1897

Db 2546 TAACTTATATGCGCGCAACCACTATGAGATTAATGATGATGCGTGTGGCAGAGACAGGA 2605

QY 1898 GTTGTGGAAATCTGCGACAGCAATTTCTTTTGTGATCTTGAATGCAATTAACATTAG 1957

Db 2606 GGCAGAGGGGTAAAGGCTGTGCTGTGACATTTAT---CTCCAGAGATGATGCAAAATATG 2662

QY 1958 CACAGCCTAGTAAAGATATGACAGATGCTCAACAGATGTTCTGATGATGGTGGAG 2017

Db 2663 CACAGATTTAGCAAGGCCCTGAGCTTCTGAGACGCAAGTCTCTGATGATCTGAAG 2722

QY 2018 AATTCCTTTAGT 2031

Db 2723 CACTTGTGATGT 2736

RESULT 15  
 US-09-923-831-42  
 ; Sequence 42, Application US/09923831  
 ; Patent No. US2002015142A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Martelange, Val,rie  
 ; APPLICANT: De Smet, Charles  
 ; APPLICANT: Boon-Faller, Thierry  
 ; TITLE OF INVENTION: TUMOR ASSOCIATED NUCLEIC ACIDS AND USES THEREFOR  
 ; FILE REFERENCE: L0461/7054  
 ; CURRENT APPLICATION NUMBER: US/09/923, 831  
 ; PRIOR FILING DATE: 2001-08-07  
 ; PRIOR APPLICATION NUMBER: 09/183, 706  
 ; PRIOR FILING DATE: 2001-10-30  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SEQ ID NO 42  
 ; LENGTH: 2365  
 ; TYPE: DNA  
 ; ORGANISM: H. sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (208)...(2151)  
 US-09-923-831-42

Query Match 7.3%; Score 158.2; DB 10; Length 2365;

Best Local Similarity 50.1%; Pred. No. 5.1e-34;  
Matches 518; Conservative 0; Mismatches 488; Indels 27; Gaps 4;

QY	902	ACAACATTTGCTAAAGCTGGTTTAACTAAGCTTACTCTGTGCAAAAAATACAGTATTCCTTA	961
DB	971	AAAAACATTTAAAAAGGAGGTTTTCMAAAGCCAAACCTATTTCAGTCACAGCATGGCCCA	1030
QY	962	TCATTTCTGACGACGAGATTTTATGCTGTGTCAACAGAGGCTGGGAAGACTGCGG	1021
DB	1031	TTGTGTTGCAAGGAATAGATCTTATAGAGTAGCCCACTGCAACGAAAGACATTTGT	1090
QY	1022	CTTTTCTCTACCAATTTTGGCTCAATGATGATGAAATTAATCTGCCAGTCGTTTAA	1081
DB	1091	GTTATTTAATGCTGATTTATTATCTGGTCTTCAAC-----CCAGCCTTAAAG	1141
QY	1082	AAGAGTTCAGAGAACCCAGAGTGTATTTATGACCACTCGAAGATTGTCAACGAGA	1141
DB	1142	GTCMAAGGAATAGACCCGGCAGTGTAGTCTAACCTCCAGGAAATAGCACTTCAG	1201
QY	1142	TTTATTTGGAAGCCAGAAAAATTTCTTTGGGACTGTGTAAAGCTGTGTATATATG	1201
DB	1202	TAGAGAGGAATGTTCGAATATTCATAT--AAAGGCTTCGAGTGTGTGTATATG	1258
QY	1202	GGGGAAACCAGCTGGGACATTCAAATTCAGCAAAATAGTACAAGGCTGTATATATATG	1261
DB	1259	GTGGTGGAATAGAGATGAACAATAGAGCTTAAAAAGGTGTAGATATCATATTTG	1318
QY	1262	CTACTCCTGGAAGCTATGATATCATAGCCAAAGAAAATTTGCTCAACAGATCA	1321
DB	1319	CAATCTCCGGGAAGTAAATGATCTGCAAAATGAGTAACTTCGTAATCTGAAGAAATTA	1378
QY	1322	AATCTAGTTTGTGATGAAGCTGATGCAATGTTGATATGGTTTGTGCCAATGA	1381
DB	1379	CTTACTGTGTTTATGATGAAGCAACAAGTGTGACATGGGATTTGAACCCAGATA	1438
QY	1382	AGAAATTAAATTTCTTGCCCAAGAAATGCCATCAAGAAACAGGCCAAACCTTATGTTCA	1441
DB	1439	TGAAGATTTTGT-----TAGATGTGCCCCAGATAGGCAACAGTATATGACCA	1486
QY	1442	GTGCAACTTTTCCAGAGAAATTCAAAGTGTGCTGCAGACTTTT---AAAGTCAATT	1498
DB	1487	GTGCTACATGGCTCATTCATTCATGCTGCTGCACAATCTTATTTGAAGAACCAATGA	1546
QY	1499	ATCTGTTTGTGCTGTGACCAAGTGGTGGAGCATGTAGAGATTCACAGACCGTTC	1558
DB	1547	TTGTCTATGTTGTGACATGGATCTAGTCTGTAAGTTCAGTGAACCAAAATATATTTG	1606
QY	1559	TCCAAGTTGGCCAGTCTCAAAAAGAGAAAGCTGTTGAATTTCTGCAAAACATAGGGG	1618
DB	1607	TAAACCAAGAGAGAAAGAAATGAGTCACATGCAAACTTTTCTACAGATATGTCATCA	1666
QY	1619	ATGAAGAACTATGCTCTTTGTGAACATAAGAAAAGCAGATTTTACTGCAACTTTTC	1678
DB	1667	CAGCAAGATCATGCTCTGCTGTTCTGAAAAGCTGTTGGGATCACTTATCAAGTAGC	1726
QY	1679	TTTGTCAAGAAAAATATCAACTCAAGTATCCATGTCATCGGGAACAGAGAGCGGG	1738
DB	1727	TAAATCTTGGAAATATATCACTAGATCTCTGCATGAGATAGAAACAGAGAGATCGGG	1786
QY	1739	AGCAAGCTCTTGAGATTTTGGCTTTGAAAAGTCCCAAGTTCCTTGTCTACTCACTAG	1798
DB	1787	AGAAGACATTAAGAACTTTTAAACAGGCAAAAGTGAAGATCTAATTGCAACTGATCTAG	1846
QY	1799	CTGCCAAGGCTGTGATATTAAGAAATGTGCAACATGTTATCAATTTTGTATCTTCTTA	1858
DB	1847	CTTCTAAGAGACTTATGATGTCATGACGTTACACATGTCATATATTTTGTACTTCCACGA	1906
QY	1859	CCATTGATGAATATGTCATGGAATGGGCTACTGGTGTGGGATATCTGSCAGAG	1918
DB	1907	ATATTTGAAGATATGTCACCGAATAGGCGCACGGAAGAGCAGAGACTGTGTTT	1966
QY	1919	CAATTTCTTTT 1931	

DB 1967 CCATTACAACTTT 1979

Search completed: June 10, 2003, 12:24:46  
Job time : 303 secs